HOUSE OF FINANCE

SAFE

WORKING PAPER SERIES

Markus Kröll - Devesh Rustagi

Shades of Dishonesty and Cheating in Informal Milk Markets in India

SAFE Working Paper No. 134

Non-Technical Summary

Markets with asymmetric information provide a fertile ground for cheating, which is a widespread problem, especially in developing countries. In order to counteract the incentives generated by asymmetric information, economists have conventionally focused on extrinsic factors, such as reputation and enforcement; the assumption being that individuals are inclined to cheat unless such factors prevent them from doing so. Yet field evidence underlines that individuals display persistent heterogeneity in cheating behavior even in the absence of extrinsic incentives. This intriguing observation together with the growing literature in economics on psychological costs of lying suggests that intrinsic motivation for honesty might also play an important role in the decision to cheat. In this paper, we examine in the field whether variation in the motivation for honesty among sellers explains differences in cheating behavior of these sellers in natural markets with asymmetric information.

A major challenge in establishing this association is to obtain a high-quality measure of motivation for honesty at the individual level. Because separating motivation for honesty from other motives is tedious while using field data, economists resort to laboratory experiments, which allow for careful separation of confounding motives. One popular experiment involves the repeated use of a random device (e.g. toss of a coin, roll of a die) and then applying statistical distributions on self-reported outcomes by the participants to classify them into honest and dishonest types. However, such a binary classification does not fully capture the heterogeneity in cheating behavior. We therefore present a novel experimental design to obtain high-quality measures of motivation for honesty at the intensive margin at the individual level. We then combine these measures with a unique objective field outcome on cheating of the same individual to investigate the importance of motivation for honesty in natural markets with asymmetric information.

Our study takes place in an important market for human health and nutrition - informal markets for buffalo milk in Delhi, India. These markets provide a unique setting to investigate our research question. We use the percentage of added water in milk as our measure of cheating behavior in the field. To obtain this outcome, we hire local assistants to pose as occasional customers and buy milk from each milkman. We then test these samples for added water using an ultrasonic milk analyzer, specifically calibrated for the Indian milk market. The milk analyzer provides us with the precise percentage of added water in milk, i.e. milliliter of water added to 1 liter of pure milk. The results reveal substantial variation in added water across milkmen.

To measure the milkmen's motivation for honesty, we conduct a behavioral experiment. Milkmen have to roll a die 40 times and are paid 2 Indian rupees (INR) for each point they report on a game sheet. The die we use is Bluetooth enabled, which transmits the actual outcomes of the rolls. This novelty allows us to contrast reported with actual outcomes and hence construct precise measures of motivation for honesty. We find considerable differences in the individuals' motivation for honesty. At the extensive margin, we find that 12.5 percent of the milkmen are classified as dishonest using statistical distributions. However, there is large variation in the degree of dishonesty even among the dishonest milkmen, which we capture via our measure at the intensive margin. This measure is the number of over-reported rolls, computed as the number of times the reported outcome exceeds the actual one. On average, milkmen over-reported in 3.63 rolls, but there is a wide variation in over-reporting, ranging from 0 to 27 rolls.

We then combine our experimental measures with the field outcome on cheating of the same milkman to test whether motivation for honesty explains the large variation that we observe in the field outcome on cheating. We find a strong and highly significant association between motivation for honesty and cheating behavior in milk markets. Dishonest milkmen add

significantly more water to milk than honest milkmen, the difference being 13 percentage points. This gap widens with the degree of dishonesty, as reflected by our measure of motivation for honesty at the intensive margin. Regardless of whether we use the extensive or the intensive margin, our measures of motivation for honesty are highly statistically significant and explain a large portion of the observed variation in the outcome. Moreover, they are also economically highly significant - dishonest milkmen earn on average INR 13,100 (~ USD 210) from adding more water, which corresponds to 7 percent of the average monthly income of honest milkmen, suggesting that intrinsic costs of lying could be large.

Our findings have important implications for both policy makers and researchers working on dishonesty. On the one hand, our study demonstrates that due to the low competition in our setting, individuals with intrinsic motivation for honesty are not driven out of actual markets and can coexist with dishonest competitors. Policies that focus on developing and promoting such honest values might thus play an important role in alleviating cheating and corruption. This also means that reliance on strong punishment for cheating might even backfire, if it crowds out intrinsic motivation for honesty. On the other hand, our study presents new challenges to the existing models of cheating. So far, economists have incorporated motivation for honesty as an internal cost that increases the incentives required for individuals to cheat. These models predict that once incentives to cheat surpass the moral costs of dishonesty, individuals would cheat maximally. We, however, do not find such a binary relationship between motivation for honesty and cheating. Our results rather suggest that there are also intermediate levels of dishonesty, which might reflect that internal costs of dishonesty are increasing in the magnitude of cheating.

Shades of Dishonesty and Cheating in Informal Milk Markets in India

Markus Kröll and Devesh Rustagi*

Abstract

Intrinsic motivation for honesty is perceived as an important determinant of large and persistent variation in cheating behavior. However, little is known about its actual role due to challenges in obtaining precise measures of motivation for honesty, as well as field outcomes on cheating. We fill these gaps using a unique setting of informal milk markets in India. A novel behavioral experiment, which combines a standard die roll task with Bluetooth technology, is used to measure motivation for honesty of milkmen at both extensive and intensive margins. We then buy milk from the same milkmen and show that cheating in the field, measured by the amount of water added to milk, widens significantly with a milkman's degree of dishonesty. Additional analyses show that conventional binary measure of motivation for honesty suffers from measurement errors, resulting in underestimation of this association.

JEL: C93, D03, O12

Keywords: Motivation for honesty, cheating, informal markets, die game milk, India.

*Goethe University Frankfurt, Faculty of Economics and Business Administration, Theodor W-Adorno Platz 4, D - 60323 Frankfurt am Main, Germany. Markus Kröll: kroell@econ.uni-frankfurt.de. Devesh Rustagi: rustagi@econ.uni-frankfurt.de. We gratefully acknowledge financial support from Sustainable Architecture for Finance in Europe (SAFE), Milkotronic Ltd. for providing us with the milk analyzer, and Ms. Maria Arbova for advising us on the machine usage. We especially thank milkmen who participated in our study, our research assistants, as well as Iqbal Malik and the NGO Vatavaran, who supported our field work in Delhi. We are also grateful to Rema Hanna and Karthik Muralidharan for helpful comments, and seminar participants at Goethe University Frankfurt, LMU Micro Workshop, NGO Conference Bad Homburg, Nova School of Business and Economics Lisbon, Stockholm School of Economics, Thurgau Institute of Economics, University of Aarhus, University of Cologne, UC Davis, UCSD Rady School of Management, University of Lausanne, University of Trier, University of Wageningen, and University of Zurich.

I. Introduction

Large and persistent variation in cheating behavior, even under asymmetry of information, poor enforcement, and weak fines, raises the question concerning motivations underlying such behavior. In the standard economics approach, individuals decide to cheat if the expected return from cheating is high and the probability of detection, as well as the severity of fines is low. This, however, stands in stark contrast to the observation that many individuals refrain from cheating despite low monitoring rates and negligible fines. Prominent examples include higher tax compliance (Andreoni et al. 1998), attendance of teachers and health workers (Chaudhury et al. 2006), and parking behavior of diplomats (Fisman and Miguel 2007). Notably, individuals differ not only in their willingness to cheat, but also in the degree of cheating: while some seize the opportunity for maximum monetary gains, others are willing to forego such gains by resorting to various levels of weak cheating (Mazar et al. 2008, Duflo et al. 2012, Pruckner and Sausgruber 2013, Fischbacher and Föllmi-Heusi 2013).[1] Thus, in addition to extrinsic incentives, intrinsic motivation for honesty, whereby individuals experience psychological disutility from dishonesty, might also play an important role in explaining cheating behavior (Gneezy 2005, Kartik 2009, Sutter 2009, Fischbacher and Föllmi-Heusi 2013).

In this paper, our objective is twofold. We first provide an innovative methodological tool to assess individual variation in motivation for honesty at both the extensive and intensive margin. Second, we investigate whether the observed variation explains differences in cheating outcomes in naturally existing milk markets in India, which are vital for human health and nutrition.

Measuring motivation for honesty is challenging. Measures that rely on observational data are likely to be of little value because of confounding with other motivations, calling for the use of behavioral experiments to carefully separate different motives. One such experiment involves self-reporting of outcomes of a random device used in private to assess dishonesty at the group level.[2] Hanna and Wang (2015) adapt this approach to measure dishonesty at the individual level by increasing the number of repetitions of the random task per individual. This adaptation allows for the use of a binary classification, whereby individuals whose self-reported outcomes exceed a chosen statistical threshold of the theoretical distribution of the random device are classified as "dishonest" and the remaining as "honest".

Motivation for honesty, however, is more subtle and complex than what the binary measure is able to capture. The binary measure is prone to measurement errors, in part

[1]As an example, Pruckner and Sausgruber (2013) find that 27 percent of the individuals pay a positive amount for newspapers at honor stands, but it is less than the designated price. Similarly, Duflo et al. (2012) find that the attendance frequencies of teachers in the comparison group varies from 8 to 24 days in their 25 visits.

[2]See for instance, Fischbacher and Föllmi-Heusi 2013, Hruschka et al. 2013, Abeler et al. 2014, Cohn et al. 2014, Dai et al. 2016.

because of the statistical test procedures, which could result in the misclassification of dishonest individuals as honest and the other way around. Moreover, it is unable to capture motivation for honesty at the intensive margin, which might further reinforce the measurement error (Mazar et al. 2008, Fischbacher and Föllmi-Heusi 2013, Pruckner and Sausgruber 2013, Rosenbaum et al. 2014).

However, little is known about the actual extent of heterogeneity in motivation for honesty, i.e. both the incidence and the degree of dishonesty, how this relates to cheating outcomes in the field, and the magnitude of the measurement error along with the subsequent bias arising from disregarding this heterogeneity in the binary measure. Resolving these gaps requires combining experimental measures of motivation for honesty with field outcomes on cheating of the same individual. But, this is demanding, as many individuals might be inaccessible to take part in an experiment or have reputational concerns over revealing dishonesty. Moreover, individuals might reduce their cheating behavior in the field during the measurement period or might even become more secretive (Banerjee et al. 2012).

We overcome these concerns by conducting our study in a setting, which allows us to obtain experimental measures of motivation for honesty, as well as an objective field outcome on cheating of the same individual. The setting is that of informal markets for buffalo milk in Delhi (India).[3] Milkmen in these markets rear buffaloes for a living and sell fresh buffalo milk directly to their customers without any intermediaries. These markets, however, are prone to asymmetric information as both ex-ante and ex-post detection of milk quality is challenging. This asymmetry of information together with the near absence of monitoring of these markets by the government and NGOs provides milkmen strong incentives to cheat by adulterating milk with water (NSMA 2011). Importantly, variation in added water cannot be explained by differences in the milk price charged to customers as milkmen collude on price in the neighborhoods they reside. The field setting thus presents a unique environment to examine the importance of motivation for honesty for cheating outcomes in the field.

We measure motivation for honesty of milkmen via a behavioral experiment that builds on the design of Hanna and Wang (2015). In the experiment, milkmen have to roll a die privately 40 times and self-report its outcome, whereby each self-reported point yields 2 Indian rupees (INR). Our experiment uses a Bluetooth enabled die, which transmits the actual outcomes of die rolls to the experimenter without the knowledge of the milkmen.[4]

[3]India is the largest producer of milk in the world (NDDB 2014). 80 percent of the milk demand in India is met by informal milk markets (Yu 2015), which are considered crucial for human health and nutrition. According to the FAO (2013), bovine milk is recognized for its potential to alleviate malnutrition and hidden hunger. This is of special relevance in India, which has high rates of malnutrition.

[4]In this aspect, our study follows other prominent studies like Andreoni (1988) and Gächter and Thöni (2005), who used such a design and withheld information to study the effect of surprise re-starts and sorting by types, respectively (see also Di Tella et al. (2015) and Maggian and Villeval (2015)). This allows us to obtain otherwise inaccessible data, which is essential to measure heterogeneity in motivation

This novel feature allows us to contrast self-reported and actual outcomes and construct refined measures of motivation for honesty that are less susceptible to the measurement error and at the same time reflect different degrees of dishonesty.

We count for every milkman the number of times the self-reported outcome exceeds the actual one and use this *number of over-reported rolls* as our main measure of motivation for honesty. Our experimental findings reveal substantial heterogeneity in the willingness to report dishonestly, with half of the 72 milkmen in our study over-reporting at least once, while the remaining 36 milkmen report honestly in all their rolls. However, there is also large variation in the degree of dishonesty: the number of over-reported rolls ranges from 0 to 27 rolls. Despite this pronounced deviation from honest reporting, the binary measure based on self-reported outcomes alone classifies only 9 milkmen as dishonest; the remaining 27 milkmen who also over-reported – albeit by a lesser degree on average – are classified as honest, resulting in sizable measurement error of 38 percent. We refer to these two groups of dishonest milkmen as "strongly dishonest" and "weakly dishonest", respectively.

We then proceed by collecting data on field outcomes on cheating. We hire assistants unknown to the milkmen to pose as potential customers and buy milk from each milkman, which is then tested for the quantity of added water using an ultrasonic milk analyzer. Adulteration of milk with water is a highly relevant measure of cheating in our setting. First, adulterated milk is sold to customers as pure milk. Second, dilution of milk with water reduces the nutritional value of milk (FAO 2013). Third, the added water is sourced from the ground, which is contaminated with harmful pollutants, especially in areas where milkmen reside (CGWB 2015). We find substantial variation in the field outcome on cheating, with added water to milk ranging from as low as 4 percent to as high as 37 percent, the average being 17.96 percent (s.d. 7.49).

Next, we investigate the association between motivation for honesty as measured by the number of over-reported rolls and cheating in milk markets measured by the percentage of added water to a liter of milk. We find a strong association between the two: the more dishonest milkmen are in the experiment, the more water they add to milk sold to the customers. A one standard deviation increase in the number of over-reported rolls (6.60) is associated with an increase in added water by nearly 3 percentage points, which is nearly one-half of the standard deviation in added water to milk. Our measure of motivation for honesty is significant at the 1-percent level in the regressions and explains up to 14 percent of the observed variation in the outcome. This association is also economically significant, as our rough estimates suggest that the additional monetary gains from dishonesty correspond to up to 7.8 percent of a milkman's monthly income. Our results hold to a powerful set of controls including the price at which the milk sample was purchased, milkmen specific socio-demographic variables, livestock related input factors, and fixed

---

for honesty and gauge the extent of measurement error in the binary measures.

effects for the time of the day the milk sample was purchased, assistants who bought milk, and the institutional environment via the dairy neighborhood where a milkman resides. We further ensure the robustness of our results to alternative measures of motivation for honesty, dropping one dairy at a time, and controlling for additional variables including bad luck.

Finally, we highlight the importance of variation in the degree of dishonesty by examining our main results separately for strongly and weakly dishonest milkmen. Our findings reveal that these two types are not just an artifact of the experiment. Weakly dishonest milkmen add significantly more water to milk than honest milkmen, but strongly dishonest milkmen cheat by adding even more water. When we contrast these estimates from those obtained using the binary measure, the downward bias turns out to be 26 percent. Despite its susceptibility to measurement errors, the binary measure is significantly associated with field outcome on cheating, suggesting that it could still prove useful as a qualitative predictor of cheating in the field. Overall, these results point not only towards heterogeneity in the willingness to cheat and different degrees of dishonesty, but also the important role it plays in explaining cheating in the field.

This paper relates to a number of strands in the literature. First, our paper contributes to a growing body of literature on the measurement of motivation for honesty using random devices. Previous studies rely exclusively on self-reported outcomes to assess motivation for honesty at the group (Fischbacher and Föllmi-Heusi 2013) or individual level (Hanna and Wang 2015). We build on these studies by presenting an innovative method which allows us to obtain refined measures at the individual level that not only capture the heterogeneity in motivation for honesty, but also mitigate the scope of measurement errors. Second, our study complements recent papers in economics that highlight the importance of motivation for honesty for field outcomes on cheating, such as selection into the public sector (Hanna and Wang 2015) and free riding on public-transport (Dai et al. 2016). Our innovative experimental design allows us to study this association in much more depth and generate important additional insights on the subtle and complex nature of motivation for honesty that the previous studies are unable to capture. Moreover, our study provides first evidence on the importance of motivation for honesty for explaining cheating in informal markets. Third, in this aspect, our paper fills an important gap in the literature on the role of motivation for honesty in the domain of informal markets with asymmetric information and thereby helps improve our understanding of the functioning of such markets (Akerlof 1970). Although motivation for honesty is hypothesized to matter, these studies only allude to this while testing for a wide variety of extrinsic factors (Dulleck 2006, Dulleck et al. 2011, Balafoutas et al. 2013). In the broader sense, our paper also connects to the literature on the role of non-monetary incentives in markets (List 2006, Dugar and Bhattacharya 2016) and the interplay of markets and morale (Bowles 1998, Shleifer 2004, Falk and Szech 2013, Bartling et al. 2015). Lastly, our paper adds to the growing literature on the role of material incentives and internalized norms in counteracting the level of cheating and corruption. While some studies have focused on the importance of monitoring (Duflo et al. 2012) and auditing (Olken 2007), others have highlighted the role of norms of corruption (Fisman and Miguel 2007) and moral reminders (Pruckner and Sausgruber 2013). Our results add to this by explicitly measuring and outlining the importance of intrinsic motivation for honesty. In this aspect, our paper can also be seen as contributing to the literature on corruption (Olken and Pande 2012, Banerjee et al. 2013).

The rest of the paper is organized as follows. Section II describes our research design including the behavioral experiment to measure motivation for honesty, our measure of the field outcome on cheating, and the sampling strategy. Section III outlines our results in three steps, beginning with patterns of motivation for honesty and the heterogeneity therein, its association with cheating in the milk market, and finally the role of the degree of dishonesty for this association including the measurement error arising from disregarding it. Section IV presents robustness checks of our results. Section V concludes with a summary of our findings.

II. Research Design

Our research design has three aims: (i) eliciting motivation for honesty at the individual level; (ii) testing the importance of motivation for honesty in explaining field outcomes on cheating; and (iii) underscoring the importance of varying degrees of dishonesty by quantifying the measurement error and the associated downward bias in the binary measure. To achieve these aims, we conduct our research in informal markets for buffalo milk in Delhi, which comprise milkmen who rear buffaloes and sell fresh unpasteurized buffalo milk directly to their customers without any intermediaries. This unique setting not only allows us to experimentally elicit milkmen's motivation for honesty but also relate these with reliable objective measures of field outcomes on water added to milk by the same milkmen. We first describe the behavioral experiment that we conducted to elicit milkmen's motivation for honesty, followed by the field outcome on cheating. The procedures that we adopted for our experiment and to collect data on the field outcome are described in Appendix B.II and A.I, respectively.

II.A. Measuring Motivation for Honesty

Measuring motivation for honesty is challenging. In the field, it is difficult to separate motivation for honesty from other confounding motives accruing from opportunities for repeated interaction, reputation formation, and as a best response to an institutional environment. Behavioral experiments allow researchers to exert control over these factors and thereby obtain clean measures of motivation for honesty. Below we describe our experimental strategy to measure motivation for honesty that builds on games of chance.[5]

Experimental strategy: Games of chance

The games of chance involve self-reporting of outcomes of random events (Fischbacher and Föllmi-Heusi 2013, Hruschka et al. 2013, Abeler et al. 2014, Cohn et al. 2014, Hanna and Wang 2015, Dai et al. 2016). In these games, individuals have to roll a die or flip a coin in private and then self-report its outcome. The payoffs depend entirely on the self-reported outcomes, providing individuals an incentive to report dishonestly by inflating the actual outcome. As there are no material gains from honest reporting in the game, any deviation from dishonesty is interpreted as reflecting internalized motivation for honesty.

Although the game is simple and easy to implement, measuring deviations from dishonesty is challenging. Because the experimenter does not observe the actual outcome of random events, dishonest reporting can only be inferred at the group level by comparing self-reported outcomes with the corresponding theoretical probability distribution of the random events (for example, Fischbacher and Föllmi-Heusi 2013). Hanna and Wang (2015) adapt this experimental strategy to measure dishonesty at the individual level by allowing for a sufficiently large number of repetitions of the random task by the same individual. Using this adaption, an individual's self-reported outcomes can now be used to assess dishonesty via a binary classification – individuals whose sum of self-reported outcomes falls below the critical threshold at a given level of significance are classified as "honest" and the remaining as "dishonest".

This binary measure, however, is potentially susceptible to measurement error for two reasons. First, some individuals are likely to be erroneously classified as "dishonest" because their self-reported outcomes surpass the critical threshold by pure chance. On the other hand, some individuals are likely to be classified as "honest" despite severe over-reporting, if the self-reported outcome falls below the critical threshold.[6] In principle, these errors can be mitigated by increasing the level of significance to reduce the number of false negatives and by increasing the number of times an individual has to roll a die to improve statistical power. Finding the appropriate number of repetitions per individual, however, is intricate: one has to not only correctly predict the effect size (the degree of over-reporting) for a given number of repetitions, but also consider pragmatic issues

[5] See Rosenbaum et al. (2014) for an overview of other experimental designs to measure motivation for honesty.

[6] For example, let us assume that an individual rolls a six-sided die twice and obtains '2' and '3' as actual outcomes, but self-reports '11' as the sum of outcomes over both rolls. If the experimenter uses the 5-percent level of significance to determine the critical threshold, he would only classify those individuals as "dishonest" who report a '12'. Thus even though the individual more than doubled the actual outcomes, he would be classified as "honest".

concerning the implementation of the experiment (e.g. amount of time, fatigue, and tediousness of the task, which increase with the number of repetitions).

Second, a simple dichotomy of "dishonest" and "honest" types is unlikely to capture fully the heterogeneity in motivation for honesty. This is corroborated by experimental evidence alluding to a wide variation in the degree of dishonesty. For example, Fischbacher and Föllmi-Heusi (2013) find that though 20 percent of the participants in their experiment over-report the outcome of a single die-roll, they do not report the payoff-maximizing outcome. Notably, this pattern does not disappear over time and is robust to a series of treatment variations. Similarly, Mazar et al. (2008) find that even though participants inflate the number of correctly solved mazes while self-reporting, they refrain from claiming to have solved the full set of mazes. This observation is not just confined to laboratory experiments. A field experiment in Austria in the context of honor stands for newspapers shows that 27 percent of the individuals pay a positive amount, but it is less than the designated price of a newspaper (Pruckner and Sausgruber 2013). Similarly, in a study by Duflo et al. (2012), attendance frequencies of teachers in the comparison group varies from 8 to 24 days over the assessment period of 25 days. Thus, when relying on the binary measure, it is hard to provide evidence for the existence let alone quantify different degrees of dishonesty, exacerbating the potential for measurement error.[7]

Experimental design

We append a die-roll task with a novel technology that allows us to record and compare the actual outcome of a random event with the one self-reported by each individual milkman. In the experiment, milkmen have to roll a six-sided die 40 times and report the outcome of each roll on a game sheet by striking out the appropriate number of 2 Indian Rupee (INR) coins.[8] For every coin struck on the game sheet, a milkmen is paid INR 2. Earnings in the game thus range from INR 80 (reporting all 1's) to INR 480 (reporting all 6's) and increase linearly in the number of reported points, providing milkmen with an incentive to over-report. We opted for this particular reporting scheme to make payoffs more salient given that the pool of participants had no prior experience in economic experiments. Instructions were neutral and did not encourage dishonesty but explicitly stated to roll the die and report the number (Appendix B.I). The responses in the post-game survey confirm that milkmen were aware of the possibility to increase their

[7] There is also a potential second order effect on measurement error. If varying degrees in dishonesty are partly driven by a motivation to increase credibility in one's report as speculated by Fischbacher and Föllmi-Heusi (2013), this would severely hamper the ability to reduce the number of false positives. In that case, individuals would just adapt their reporting behavior accordingly, once the number of reported outcomes is increased.

[8] We followed the design by Hanna and Wang (2015) who compute the necessary number of repetitions to be 37 to achieve a power of 80 percent at a significance level of 5-percent based on the (implicit) assumption that an individuals' reporting behavior in each roll is an independent draw from the aggregate reporting behavior observed in the study by Fischbacher and Föllmi-Heusi (2013).

payoffs by over-reporting the actual outcomes.

We conducted the experiment with a Bluetooth enabled die, which transmits the actual outcome of each roll to a hidden smartphone of one of the authors.[9] Thus, we are able to observe the number of times a milkman over-reported, as well as the magnitude of over-reporting. In addition, this adaptation allows us to compare our results with self-reported measures and quantify the measurement error.

While conducting the experiment, we took great care to ensure that every milkman understood the game, carried out the experiment as outlined before (e.g. rolled the die) and that contagion across participants was minimized. After all milkmen took part in the experiment, they were invited to fill a post-game survey in private. Upon the completion of the survey, milkmen were paid their earnings from the experiment. On average, milkmen earned INR 495 ($\sim$ USD 8) including a show-up fee of INR 200.

II.B. Measuring Field Outcome on Cheating

Field setting

The informal milk markets in Delhi are prone to asymmetric information. Milkmen milk buffaloes privately in the absence of customers, as well as other milkmen inside a stable housing several buffaloes, making ex-ante detection of milk quality challenging. Similarly, ex post detection of adulteration is severely impeded by the high fat content of buffalo milk (9 percent), which is twice the fat content of cow milk. While there are simple milk testing procedures that can be implemented at home, results from a field experiment that we conducted show that these tests are unreliable (see Appendix A.III).[10] Analysis of milk samples in formal food testing laboratories is beyond the reach of customers, because it is very expensive and restricted to large industrial orders.

A peculiar feature of these informal markets is that individual farms are found in geographical clusters, each labeled a "dairy", where animal husbandry is permitted under the zoning laws of Delhi. There is limited competition within and between dairies. Within a dairy, the close-knit communities of milkmen form and enforce a cartel to collude on the price of milk and to dissuade fellow milkmen from luring customers away from each other. Competition across dairies is mostly lacking because of large distances separating them. Customers usually come from within a radius of a maximum of five kilometers surrounding a dairy, but the average distance between dairies is much more than that.

[9] We also conducted a control experiment in which half of the participants played the game with a Bluetooth die and the other half with a standard die to show that our measures are not affected by the use of a Bluetooth die itself. We find that there is no difference in the reporting behavior across the two groups of participants (see Appendix A.IV).

[10] We dilute pure buffalo milk with different levels of water ourselves and then invite milkmen to rate each sample on milk quality in an incentivized tournament. Our results reveal that milkmen are unable to distinguish pure from highly diluted milk.

Moreover, competition with the formal milk sector is low due to differences in the milk sold.[11]

This asymmetry of information together with the near absence of monitoring of milk quality sold in these markets by the government or other agencies including NGOs provides milkmen a strong incentive to cheat, which is further exacerbated by limited competition. In fact, a study conducted by the Food Safety and Standards Association of India (FSSAI) found that 70 percent of the 71 milk samples from Delhi were adulterated with water (NSMA 2011).[12]

Field outcome on cheating

Following the results of the FSSAI-study (NSMA 2011), we use added water in buffalo milk as our measure of field outcome on cheating. We assess this using an ultrasonic milk analyzer, which calculates the percentage of added water in a liter of milk based on the freezing point of milk.[13]

Added water is a highly relevant measure of cheating in our field setting for three reasons. First, dilution with water reduces the nutritional value of milk and thus its potential to alleviate malnutrition and hidden hunger (FAO 2013). This is also corroborated by our data. We find a strong negative correlation between the percentage of added water and the amount of solids-not-fat (i.e. the micronutrients including calcium, phosphorus, vitamins, etc.) and protein in milk (r>0.98, p-value = 0.000). Second, because the added water is sourced from the ground, which is contaminated from untreated sewage in areas where milkmen reside (CGWB 2013), it additionally exposes customers to long-term health risks. Hence, adding water is not just a sheer redistribution of income, but harms society. Third and importantly, the practice of adding water is also not an implicit arrangement between milkmen and their customers to provide a lower-quality product in exchange for a more affordable price. The buffalo milk in informal markets costs on average INR 57 per liter and is more expensive than a liter of pasteurized buffalo milk in the formal market (INR 50). A household survey in which we asked milkmen to state the

[11] The formal sector comprises government-backed producers, cooperatives, as well as private companies, such as Mother Dairy, Amul, and Parag Dairy respectively, and mostly sells treated, pasteurized milk. Additionally, it trades almost exclusively in cow milk, which contains less fat and other nutrients.

[12] This pattern is not restricted to Delhi alone but is observed across India. Of the 1791 samples collected from all over India, 70 percent tested positive for adulteration with water (NSMA 2011).

[13] The base freezing point of buffalo milk is a physiological constant that lies below the freezing point of water. However, as more water is added to milk, its freezing point moves closer to that of water (Advanced Instruments 1995). The precise formula used by the machine is:

$$Added\ Water\ = \frac{Freezing\ Point_{Base} - Freezing\ Point_{Sample}}{Freezing\ Point_{Base}} \cdot 100[\%].$$

The correct calibration of the base freezing point is crucial to obtain reliable measures of added water in milk. The producer of the machine, Milkotronic Ltd., maintains a large database on buffalo milk in India and calibrated the base freezing point accordingly. For more details, see `http://www.milkotronic.com/pdfs/Lactoscan_SA_Eng.pdf`.

milk quality their customers demand and that sold by them confirms this. All milkmen stated that customers demand pure, unadulterated buffalo milk and denied adding any water to milk. Many milkmen stated further that if customers wanted diluted milk they could add water themselves. Lastly, our survey data also reveals that fluctuation in milk demand is very low and occurs at the most once a month.

Figure 1 shows the distribution of added water in milk. On average, 17.96 percent (s.d. 7.49) of the milk sold by milkmen is actually added water. However, there is a large variation, ranging from as low as 4 percent to as high as 37 percent.

We took several steps to ensure the reliability of our field measure on cheating. First, we did not buy milk ourselves but instead hired several local assistants unknown to the milkmen and instructed them to pose as occasional customers and buy a liter of milk from specific milkmen in our sample. Because milkmen had no opportunity to find out that the same team is behind both the behavioral experiment and milk purchase, they could not alter the milk quality sold to our assistants and treated them like other customers. Second, even though informal milk markets are largely based on repeat interaction but our assessment of the field outcome is based on milk sold to occasional customers, it does not compromise the generalizability of our findings. Results from household and community surveys reveal that milkmen do not discriminate between occasional and repeat customers either in quality or price. Both are sold milk from a large container, which a milkman uses to store the entire milk for sale. Third, to add credibility to the representativeness of our sample, we bought milk from 63 milkmen who did not take part in the experiment to show that there is no difference in added water between these and milkmen who took part in the experiment (p-value = 0.58).[14] Fourth, we verify in a pilot study that results on added water obtained using the machine are strongly correlated with an independent analysis of the same milk by a professional food-testing laboratory ($r = 0.93$, see Appendix A.I). In addition, we were worried that outcomes based exclusively on added water might understate cheating in the field, if milkmen add other adulterants to milk besides water (NSMA 2011). Therefore, we additionally requested the laboratory to test each sample for other potential adulterants. The results are discussed in Appendix A.I and confirm that added water is the only adulterant. Fifth, we also internally validate our field measure through the purchase of a second milk sample from 50 percent of the milkmen in our study. We find a strong and highly significant correlation in added water between the two samples for this sub-set (p-value = 0.000).

II.C. Sample Construction and Characteristics

Our sampling strategy focuses on small milkmen who reside within the dairy neighborhoods and run the day to day operations themselves. Such a strategy allows us to map

[14]This result also holds when we control for the institutional setting using dairy fixed effects.

tightly the association between motivation for honesty and field outcomes on cheating. Accordingly, we restrict our sample to milkmen from six dairies, of which five are from Delhi and the sixth one is from the adjoining city of Sahibabad, which is located in the state of Uttar Pradesh and borders Delhi on the east.[15] The minimum distance between any two dairies in our sample is 21 km.

There are roughly 160 milkmen residing in these six dairies who meet our sampling criteria. Even though this is not a large population, it was not possible to include all of them in our study. The use of the Bluetooth die meant that every milkman required our personal attention. As a result, we were not able to run the experiment with multiple milkmen at the same time. Moreover, we had to conduct the experiment within the limited time that the daily routines of the milkmen permit; this further puts a limit on the number of milkmen that we could invite to take part in the experiment.[16]

Led by these considerations, we have 84 milkmen participating in our experiment, implying a representation rate of 53 percent. Milkmen were recruited with the help of a "community mobilizer", a local milkman who assisted us in this process. Despite our best efforts, we could not locate the farm of nine milkmen and for another three milkmen, the Bluetooth outcome is missing.[17] Excluding these 12 observations leaves us with a final sample of 72 milkmen.

Panel C, Table 1 reports summary statistics on milkmen in our sample. On average, a milkman is 34 years old, has 10 years of education, low self-reported proficiency in English, and owns 19 buffaloes. Nearly 50 percent of the milkmen belong to *Gujjar* or *Yadav* caste. 79 percent own at least one buffalo which had begun yielding milk in the three months prior to our study.

III. Results

We present our results in three steps, which are aligned with our three aims. We begin with results from the behavioral experiment to outline our measures of motivation for

[15]While most dairies largely comprise such small milkmen, there are also large farms that house up to several hundred buffaloes. In these farms, daily activities are exclusively carried out by hired labor. The owners reside outside the dairy and take on the role of supervisors visiting the farm only at the time of milk sale to collect revenues. Accordingly, the dairies in our sample were selected on the basis of information gathered via a baseline survey conducted on our behalf by a local NGO. In this survey, we collected information on a variety of dairy-specific topics, e.g. the size of the dairy, size of the farms, labor employment, farm management, and number of livestock. We excluded five dairies from Delhi; three of these comprise exclusively of large milkmen, while the other two are very small and unviable for the study.

[16]In the four experimental pilot studies that we conducted, a session with 15 participants lasted up to three hours including the post-experimental survey and payments. Milkmen sell milk from 6 am to 10 am and then again from 3 pm to 6 pm. Extra time is devoted to feeding and washing buffaloes before milking.

[17]Our results do not change if we include these three milkmen in our sample. We also verify that these 12 milkmen do not differ from the rest either in their self-reported behavior in the experiment or their socio-demographic characteristics (Table A1, Appendix A.II)

honesty that rely on actual outcomes obtained using the Bluetooth die. We then contrast these with the results obtained from using conventional measures that rely only on self-reported outcomes. Next, we combine our experimental measures with our field outcome on cheating to underscore the importance of motivation for honesty in informal milk markets that are vital for human health. Lastly, we highlight specifically the importance of varying degrees of dishonesty in such markets and assess the downward bias arising from measurement error in conventional measures.

III.A. Motivation for Honesty

We now describe our experimental findings, beginning with the measures, followed by patterns and correlates of motivation for honesty.

Measures of motivation for honesty

Following the literature, we interpret honest reporting of outcomes of die rolls in the experiment as reflecting an individual's internalized motivation for honesty, especially because there are no material or other strategic gains from doing so (Fischbacher and Föllmi-Heusi 2013, Hanna and Wang 2015). In contrast to the conventional measures that rely only on self-reported outcomes, our novel Bluetooth-based measures are based on a comparison of self-reported and actual outcomes, and thus are continuous by construction. We describe these measures in detail below.

Bluetooth-based measures. – We develop four different measures, which differ in their informational content and their sensitivity to the randomness of the die-roll task. The simplest measure, *number of over-reported rolls*, treats the outcome of each die roll as a binary event and counts the incidents over all die rolls in which the self-reported outcome exceeds the actual outcome. A related measure additionally considers the magnitude of over-reporting by using the *sum of added points* over all die rolls. These two measures may be sensitive to the randomness of the task itself, the latter more so than the former.[18] To mitigate this and also account for "bad luck" in the experiment, we construct two additional measures, which express the number of over-reported rolls and the sum of added points in relative terms. To construct the *share of over-reported rolls*, we divide the number of over-reported rolls by the number of recorded rolls in which an individual did not obtain a '6'. Similarly, for the *share of added points*, we calculate the ratio between the sum of added points and the maximum number of points a milkman could have added given his realizations in the recorded die rolls. One caveat of the Bluetooth technology is

[18]While the number of points an individual can add in a single die roll always depends on its actual outcome, the decision whether to over-report or not is only restricted by the actual outcome, when a '6' is realized. When the actual outcome is a '1', a milkman can add 0, 1, 2, 3, 4, or 5 points, but he can only add up to two points for a '4' and cannot add any points when the actual outcome is a '6'.

that due to temporary connectivity deficiencies, the Bluetooth die did not transmit the actual outcome for nearly 10 percent of the die rolls. We refer to such rolls as *missed rolls* in the remainder of the text. While the relative measures already take this caveat into consideration, the other two do not. Consequently, while conducting robustness checks (c.f. section IV), we show that our main results are not affected by missed rolls. For the ease of interpretation, we use as our main measure the number of over-reported rolls, and the remaining three as alternative measures. The summary statistics on these measures is reported in Panel B of Table 1.

Conventional measures. – In line with the previous studies, we use the binary measure, whereby the normal distribution is used to define a critical threshold for the number of self-reported points that are sufficiently unlikely to be accumulated over 40 die rolls. We then classify all those milkmen as "dishonest" whose self-reported points fall above that threshold, and the others as "honest". At the 1-percent level of significance, this critical threshold is 166 points. The corresponding number at the 5-percent level of significance is 158 points. In addition, following Hanna and Wang (2015), we also consider a measure based on the *sum of self-reported points* over all the rolls. The summary statistics on these variables is reported in Panel B, Table 1.

Patterns of motivation for honesty

Figure 2 plots the sum of self-reported outcomes against the number of over-reported rolls at the individual level. Two patterns are noteworthy.

First, there is large heterogeneity at both the extensive and intensive margin. 36 out of 72 milkmen (50 percent) reported honestly in all of their recorded rolls (circles), but the remaining 50 percent reported dishonestly (diamonds). The summary statistics reported in Panel B of Table 1 show that, on average, milkmen over-reported in 3.6 rolls and added 7.5 points in total. The relative measures reveal that milkmen over-reported in nearly 11 percent of the rolls in which they did not obtain a six and added nine percent of all the points they could have possibly added. However, there is also considerable variation in the degree of dishonesty, which varies from over-reporting in 1 to 27 rolls by 1 to 74 points.

Second, conventional measures are prone to measurement errors. Although 36 milkmen actually reported dishonestly, the binary measure identifies only 9 of these milkmen (black diamonds) as dishonest using the critical threshold at the 1-percent level of significance (solid line). The remaining 27 dishonest milkmen (white and gray diamonds) along with the 36 honest milkmen (circles) are bunched together as honest. This equates to a Type-II-Error of 37.5 percent. This concern is not alleviated when we lower the level of significance to 5-percent to determine the critical threshold (dotted line), but rather worsens. While three of the 27 milkmen who were previously misclassified as honest are now correctly classified as dishonest (gray diamonds), four honest milkmen are now misclassified as dishonest (gray circles), thereby introducing Type-I-Error and increasing slightly the overall measurement error to 38.9 percent.[19] Similarly, the measure based on the sum of self-reported outcomes displays nine milkmen considered dishonest under the binary measure as having the strongest degree of dishonesty. However, it is evident from the number of over-reported rolls that seven milkmen considered honest under the binary measure actually display the same or even stronger degree of dishonesty.

Together, these results suggest large heterogeneity in motivation for honesty from the willingness to cheat, as well as the degree of cheating, which we now examine in detail. Based on the patterns described above, we identify three distinct types of milkmen and then test whether these types differ significantly in their dishonesty in the experiment:

(i) Strongly dishonest: 9 milkmen who are classified as dishonest by both the Bluetooth and the binary measure.

(ii) Weakly dishonest: 27 milkmen who are classified as dishonest under the Bluetooth measure but honest under the conventional binary measure.

(iii) Honest: 36 milkmen who are considered honest under both Bluetooth and conventional measure.

Table 2 presents results from a regression of the number of over-reported rolls on a dummy for strongly dishonest and weakly dishonest milkmen each, with honest milkmen serving as a benchmark category.[20] Column 1 is without controls, column 2 controls for milkmen specific characteristics, and column 3 additionally includes dairy fixed effects. The results reveal that weakly dishonest milkmen over-reported on average by nearly four more rolls than honest milkmen, for whom the number of over-reported rolls is zero by definition; the difference being highly significant (p-value < 0.000). Similarly, strongly dishonest milkmen over-reported by roughly 17 more rolls on average, which is also highly significant (p-value < 0.000). Importantly, the strongly and weakly dishonest types also

[19] While Hanna and Wang (2015) calculated the number of repetitions to achieve a statistical power of 80 percent at a significance level of 5 percent and thus bound the share of falsely classified individuals between 5 and 20 percent, the actual measurement error might still be much larger for two reasons. First, their power analysis is based on the (implicit) assumption that an individual's reporting behavior in each roll is an independent draw from the aggregate reporting behavior observed in the study by Fischbacher and Föllmi-Heusi (2013). This aggregate reporting behavior, however, is the average over many different types, ranging from honest and weakly dishonest to maximally dishonest individuals. It seems doubtful that every dishonest individual combines within himself this entire spectrum of behavior. Moreover, as noted before, if varying degrees in dishonesty are partly driven by the motivation to increase credibility in one's reporting (Fischbacher and Föllmi-Heusi 2013), individuals might adjust their reporting behavior according to the number of repetitions.

[20] Note that this analysis is not tautological. Although for the weakly dishonest milkmen, the sum of self-reported points is by definition lower than that of strongly dishonest milkmen, it is not clear if the same applies for the number of over-reported rolls. As an example, seven weakly dishonest milkmen over-reported more or the same as some of the strongly dishonest milkmen.

differ significantly from each other in their over-reporting (p-value < 0.000). Columns 4-6 show that our results hold when we use alternative measures of motivation for honesty including the sum of added points, share of over-reported rolls, and share of added points. These results confirm large heterogeneity in motivation for honesty: individuals differ not only in their willingness to cheat but also in their degree of dishonesty.

Correlates of motivation for honesty

Finally, we examine the correlates of motivation for honesty as measured by the number of over-reported rolls in Table A.2 of Appendix A.V. Motivation for honesty is uncorrelated with control variables included in the main regression, as well as additional control variables, which we introduce while conducting robustness checks; the only exception is the variable family size (p-value $= 0.076$). We get similar results while using a regresion framework in which we include all of these variables.

III.B. Motivation for Honesty and Cheating in Milk Markets

We now set out to investigate the role of motivation for honesty for cheating outcomes in milk markets. We present results using the Bluetooth-based measures and relegate the robustness of these findings to section IV. Results using the conventional measures are presented in the next sub-section. Our hypothesis is that cheating in milk markets as measured by the percentage of added water in milk increases with the degree of dishonesty in the experiment. We estimate the association econometrically using the following OLS specification:

$$Y_{id} = \beta_0 + \beta_1 Dishonesty_{id} + \beta_2 Price_{id} + \beta_3 X_{id} + \beta_4 L_{id} + \alpha_t + \alpha_d + \alpha_a + \epsilon_{id} \qquad (1)$$

Y_{id} is the field outcome on cheating measured as added water in milk in percent by milkman i from dairy d. $Price$ is the amount paid for the milk sample. X and L are vectors of milkmen's socio-demographic and livestock specific input factors, which are described in Panel C of Table 1. α_t, α_d, and α_a are fixed-effects for the time of day the milk sample was bought, the dairy where a milkman resides, and the assistant who bought the milk. The variable *Dishonesty* is measured as the number of over-reported rolls and captures the effect of motivation for honesty on cheating in milk markets. This choice, however, has no bearing on our main results, which we are able to replicate with all other alternative measures as well (see section IV).

Empirical strategy

A key concern in estimating equation 1 is that besides motivation for honesty, several other factors might affect a milkman's decision to cheat, ranging from socio-demographic factors (Armantier and Boly 2011, Abeler et al. 2014) and inputs to milk production to disparities in the institutional environment (Banerjee et al. 2012) and assistant specific factors (Balafoutas et al. 2013). Although we show in Table A.2 of Appendix A.V that motivation for honesty is uncorrelated with control variables, we take several steps to alleviate these concerns even further. For instance, in all our estimates we control for important socio-demographic covariates of milkmen including age, education, English proficiency, and caste. These are expected to capture differences in cheating behavior in the field that may be due to, for instance, differences in education or caste specific norms. Moreover, while conducting robustness checks, we additionally control for outside option, religiosity, and motivation for honesty of the nearest neighbor.

To ensure that our results are not capturing livestock related inputs, we control for herd size and lactation period of buffaloes in our main specification. While buffalo herd size is a proxy for overall milk output, lactation period influences fat levels in milk, which could eventually affect the level of added water. In the first three months of the lactation period, buffalo milk contains less fat and is less dense (FAO 2013); this could discourage the addition of further water to milk. In addition, we conduct robustness checks by considering proxies for buffalo breed and buffalo feeding costs.

We also ensure that weather specific and institutional differences are not driving our results. We bought milk from milkmen within a dairy on the same day and from all milkmen during the third week of December. This together with fixed effects for the time of the day we bought milk ensures that factors affecting milk production, such as temperature and humidity are already controlled for (Marai and Haeeb 2010). Moreover, we include dairy fixed effects to account for time-invariant factors across dairies, such as social norms.

We take steps to show that our results are not due to assistant specific factors either. All assistants are male in their mid-twenties, have high school education, and reside in the same neighborhood which is located outside the catchment area of the dairies in our study. We also include fixed-effects for the assistants in our specification.

Results

Table 3 presents results on the association between motivation for honesty and cheating in milk markets. Column 1 is without any additional controls and shows that dishonesty has a positive coefficient, which is significant at the 10-percent level. This implies that the percentage of added water in milk increases with the degree of dishonesty. The coefficient remains robust in magnitude once we control for the price of the milk sample in column 2, but its standard error declines, such that it is now significant at the 5-percent level. The coefficient on price is negative and significant at the 1-percent level. These results suggest that motivation for honesty plays a role over and above that of price.

We next introduce milkmen and livestock specific controls in column 3. The coefficient on dishonesty rises slightly but is now much more precisely estimated and is significant at the 1-percent level. As expected, the lactation dummy has a negative coefficient, which is significant at the 5-percent level. When we introduce our set of fixed effects for the time of the day the milk was bought, dairies, and assistants in column 4, the coefficient on dishonesty rises strongly, but its standard error remains nearly the same.[21] These results point out that milkmen add, on average, 0.44 percentage points more water to milk per over-reported roll. Put differently, one standard deviation increase in the number of over-reported rolls (6.60) is associated with a rise in added water in milk by 2.92 percentage points, which is one-sixth of the mean level of added water. Motivation for honesty has the largest effect of all the variables. Moreover, it explains 14 percent of the variation in the outcome, which is also the largest of all covariates that together account for 32 percent.

Our estimates are also economically relevant. Given that the water which is added to milk is sourced from the ground and thus free, the increase in adulteration of milk associated with a one-standard deviation increase in the number of over-reported rolls would yield milkmen higher profits between INR 1.81 and INR 4.14 per liter of milk, the average gain being INR 2.65. Over the course of a month, depending on the size of the dairy and the cheating behavior of a milkman, the additional gains would range from INR 857 to INR 39,068, which corresponds to up to 7.8 percent of a milkman's monthly income.[22] This suggests that intrinsic costs of dishonesty could be large.

Although it is beyond the scope of this study to show the negative consequences of dishonesty on human health and nutrition via the addition of water to milk, we do highlight its negative association with the amount of protein and solids-not-fat in milk, the latter containing the micronutrients. The results reported in columns 4-5 of Table 3 show that a one-standard deviation increase in the number of over-reported rolls leads to a fall in protein by 0.19 percentage points and in micronutrients by 0.41 percentage points. Given that the average levels of these nutrients in milk is 7.21 and 3.25 percent, these losses are not trivial.

Together, these results highlight not only a strong association between motivation for honesty and cheating outcomes in milk markets, but also its economic importance.

[21] Notice that the coefficient on price declines both in absolute magnitude and significance especially when fixed effects are introduced. This is because due to collusion, the variation in price is mostly between and not within dairies.

[22] To compute the additional profits from adding water, we calculate the additional revenue made with 1 liter of pure buffalo milk by increasing the level of added water in 1 liter of milk sold by 2.92 percentage points. We therefore divide the price charged for one liter of milk by a given milkman by its share of pure milk (1 - added water in percent) and compare this to the quotient of the price and its share of pure milk minus the additional 2.92 percentage points. To compute the monthly profit, we use the average number of lactating buffaloes (13.20) and assume an average milk yield per buffalo per day of 10 liters, which is based on data obtained from the household survey.

III.C. Degrees of Dishonesty and Measurement Error

We now specifically underscore the importance of varying degrees of dishonesty by showing that weakly dishonest milkmen are not just an artifact of the experiment, but also differ in their cheating outcomes in milk markets. We then compare our results with those obtained from conventional measures that are either unable to detect weak dishonesty or detect it with error. This additional step allows us to quantify the bias arising from disregarding such heterogeneity in motivation for honesty.

Degrees of dishonesty and cheating in milk markets

We modify equation 1 and estimate the association separately for strongly dishonest and weakly dishonest milkmen with respect to honest milkmen as a benchmark:

$$\begin{aligned} Y_{id} = \delta_0 + \delta_1 Strong\ Dishonesty_{id} + \delta_2 Weak\ Dishonesty_{id} \\ + \delta_3 Price_{id} + \delta_4 X_{id} + \delta_5 L_{id} + \alpha_t + \alpha_d + \alpha_a + \mu_{id} \end{aligned} \tag{2}$$

where the variables *Strong Dishonesty* and *Weak Dishonesty* capture separately the effect of strongly and weakly dishonest milkmen on cheating outcomes in milk markets. We measure strong and weak dishonesty both at the intensive and the extensive margin. While at the intensive margin we use the variable number of over-reported rolls separately for each type, at the extensive margin we simply use a dummy variable for each type. Given our experimental results that weakly dishonest milkmen differ significantly from both strongly dishonest and honest milkmen in their dishonesty, we hypothesize $\delta_1 > \delta_2 > 0$.

Columns 1-3 in Table 4 report the results with a full set of controls and fixed effects. Regardless of the measure, both strongly and weakly dishonest milkmen add significantly more water to milk than honest milkmen. According to the estimates in column 1, while a one-standard deviation increase in over-reporting by weakly dishonest milkmen is associated with a rise in added water by 0.8 percentage points, the corresponding estimate for strongly dishonest milkmen is 3.75 percentage points. We get similar results when we use the extensive margin defined at the 1-percent level of significance in column 2. Weakly dishonest milkmen add on average 3.14 percentage points more water to milk than honest milkmen, but the difference is much larger for strongly dishonest milkmen, who add on average 9.57 percentage points more water. Estimates reported in column 3 use the extensive margin at the 5-percent level. While the coefficient on weakly dishonest milkmen declines only slightly, that on strongly dishonest milkmen declines by half; nevertheless both remain statistically significant. This is because of the introduction of Type-I-Error (c.f. III A). When we correct for this error (results not shown), the coefficient on strongly dishonest milkmen rises to 7.841 (s.e. 2.170) and that on weakly dishonest milkmen to 3.060 (s.e. 1.461), which are comparable to the estimates in column 2. These results set to confirm that different degrees of dishonesty captured by our refined Bluetooth-based measure play an important role in explaining cheating outcomes in the field.

Measurement error and downward bias

Given that the cheating behavior of weakly dishonest milkmen differs significantly from strongly dishonest and honest milkmen, estimates from conventional measures that pool weakly dishonest and honest milkmen into a single category are expected to be biased downward.[23] To quantify this bias, we estimate equation 1 using the conventional binary measure that defines the critical threshold at the 1-percent level of significance. Column 4 in Table 4 presents the result and shows that the coefficient on strongly dishonest milkmen is 8.01 (s.e. 2.35) and is significant at the 1-percent level. When we compare this with the coefficient in column 2 (9.57), the downward bias turns out to be 16.4 percent. The bias gets even larger if we relax the level of significance to define the critical threshold to 5-percent in column 5. Now, the coefficient on strongly dishonest milkmen is even smaller in magnitude (3.587, s.e. 1.937) relative to the coefficient reported in column 3 (4.869, s.e. 2.163), suggesting a downward bias of over 26 percent.

The measurement error is not mitigated when we use the conventional measure based on the sum of self-reported outcomes. As mentioned in section III A, this measure is noisy due to the randomness of the die-roll task. This notwithstanding we find that the coefficient on the self-reported sum is positive (0.109, s.e. 0.039) and significant at the 1-percent level (column 6).[24] However, once we account for the randomness of the die rolls in column 7 by additionally controlling for the actual outcome of the die rolls, the coefficient rises to 0.144 (0.039) without a corresponding rise in its standard error. This implies a downward bias of 24.3 percent.

Overall, our findings suggest that heterogeneity in motivation for honesty, which can only be fully unveiled when comparing self-reported to actual outcomes, plays an important role for field outcomes on cheating. Thus, relying purely on self-reported outcomes for the assessment of motivation for honesty underestimates its association with cheating in the field. Despite this drawback, conventional measures turn out to be useful as qualitative predictors of cheating in the field.

IV. Robustness Checks

We now present results from a variety of robustness tests, which corroborate our main findings reported in Table 3. We start by using alternative measures of motivation for honesty, followed by the inclusion of additional control variables, dropping influential observations, and clustering of standard errors at the dairy level. Finally, we take several

[23]This is equivalent to the estimation of equation 2 with δ_2 set equal to 0.

[24]Hanna and Wang (2015), on the other hand, find no significant effect.

steps to address the potential error in our measure of motivation for honesty due to missed recordings of the Bluetooth die.

Alternative measures

Columns 1-3 in Table 5 report results from the estimation of equation 1 using alternative measures of dishonesty. In column 1, we use the sum of added points which allows us to go beyond the incidence and take the magnitude of over-reporting into consideration. In column 2, we express the number of over-reported rolls in relative terms as the share of over-reported rolls, which allows us to factor out the randomness of the die roll task. In column 3, we combine both of the above into a single variable by using the share of added points. Irrespective of the measure we employ, the association between motivation for honesty and cheating in milk markets remains highly statistically significant (p-value $<$ 0.001). Notably, standardized coefficients reported in the last row of the table show that estimates obtained from different measures are comparable to each other and fall between 2.5 and 2.9.

Additional control variables

We next test the robustness of our results to the inclusion of additional control variables. The results are reported in Table 6. While column 1 reports the coefficient on the additional variable, column 2 tracks changes in the coefficient on the dishonesty variable from the inclusion of the corresponding additional variable.

We begin by addressing the potential concern that milkmen add varying quantities of water to increase their total output because they have poorer milk yielding buffalo breeds. Although adding water for this reason is still cheating, we nonetheless address this concern by controlling for the number of buffaloes a milkman has from the state of Haryana, which the milkmen reported in the surveys to be of higher yielding variety than buffaloes from other states. A similar concern may arise from differences in the quality of fodder fed by milkmen to their buffaloes. It is difficult to obtain precise measures of fodder quality, so we use monthly expenditure on fodder as a proxy, the data for which was collected using a household survey. While the coefficient on buffalo breed is indeed negative and statistically significant,[25] the coefficient on feeding costs is not significantly different from zero. Importantly, the inclusion of these variables does not lead to any major changes in the magnitude or the significance of the coefficient on our dishonesty variable.

We next include one by one the following additional controls. We use duration of stay within a dairy to account for learning effects regarding the returns from cheating.

[25] We are unable to include this variable in the main specification because we could not collect data for seven milkmen despite repeated visits. All our results hold even when we drop these seven observations.

To account for the importance of home environment as a predictor of dishonest behavior (Fisman and Miguel 2007), we introduce a dummy variable for the state of origin of the milkman.[26] Furthermore, milkmen might vary in their cheating behavior due to differences in outside options. This data is difficult to obtain in our setting, so we proxy for outside option using a dummy variable, which takes the value of 1 if a milkman himself or members from his household have a part-time or full-time job outside the dairy sector; otherwise, the value is 0. Following Armantier and Boly (2011), who identify religiosity as a micro-determinant of dishonesty, we control for the average number of visits to a temple or a mosque in a month.[27] Another important factor behind differences in cheating could be family size because milkmen with larger families have higher living expenses. Lastly, a milkman might cheat because his neighbors also cheat. Even though milkmen add water to milk in private and thus cannot observe the practices of fellow milkmen, we include the number of over-reported rolls of the nearest neighbor as an additional control. As column 1 clearly shows none of the coefficients of these additional control variables are statistically significant. In contrast, the coefficient on dishonesty remains remarkably stable in magnitude (around 0.43) and significance (p-value < 0.001) all throughout the column 2. All of these results hold even when we use alternative measures of motivation for honesty (see Table A.3 in Appendix A.VI).

Dropping observations and clustering

We next proceed to show that our results are also not due to a specific dairy or specific individuals. Column 1, Table A.4 in Appendix A.VI shows a change in the coefficient on the dishonesty variable resulting from dropping one dairy at a time. Overall, we do not find much fluctuation in the coefficient, which remains significant at the 1-percent level and above 0.40 in magnitude in five out of six cases. Although the coefficient declines in magnitude to 0.307 (s.e. 0.141) when we drop dairy 5, it remains significant at the 5-percent level. Moreover, our estimated relationship is also robust to the removal of six influential milkmen found in four dairies either individually or altogether.[28] All of these results hold when we use alternative measures of motivation for honesty .

We also analyze the role of potential spatial correlation by clustering standard errors at the dairy level. We use wild bootstrap procedures to account for the small number of clusters (Cameron et al. 2008). Also with clustered standard errors, the coefficient on dishonesty remains highly significant (p-value $= 0.018$).

[26] In our sample, milkmen hail from four different states of India: Uttar Pradesh (36 milkmen), Haryana (27), Rajasthan (5), and Delhi (3). The dummy that we introduce takes the value of 1 if the home state is Uttar Pradesh, otherwise zero.

[27] 12.5 percent of the milkmen in our sample are Muslims. When we control for the religious confession using a dummy variable, our results do not change.

[28] Influential observations are identified using DFITS, which classifies observations as influential, if the difference in fitted values with and without the i-th observation is larger than $2 \cdot \sqrt{k/N}$, where k is the number of parameters and N is the sample size.

Bluetooth misses

One remaining concern is that of missed Bluetooth recordings in the experiment. While our measures of dishonesty, such as share of over-reported rolls and share of added points account for this, our results could still be biased if missed rolls represent a special from of dishonest behavior that is correlated with field outcomes on cheating. Therefore, we conduct a battery of robustness checks, the results of which are reported in Panel A, Table A5 of Appendix A.VI. In column 1, we present results from a regression in which each observation is weighted by the number of recorded rolls. In column 2, we directly control for the number of missed rolls. In column 3, we exclude observations for which we missed ten or more recordings. None of this has any major implications for our findings; the coefficient on number of over-reported rolls is always close to 0.45 and remains highly significant at the 1-percent level. As a stronger robustness check, when we assume in column 4 that milkmen always over-reported in all the missed rolls unless they reported a '1', the coefficient on dishonesty drops in magnitude to 0.28 (s.e. 0.087), but remains highly significant at the 1-percent level. In column 5, we weight the number of over-reported rolls linearly by the number of missed rolls, i.e. we assume that a milkman's reporting behavior does not differ in recorded and non-recorded rolls. Lastly, in column 6, we use the actual outcome in all recorded rolls in our sample to calculate the average share of rolls in which over-reporting occurred conditional on each reported outcome. We then replace non-recorded rolls with these averages.[29] As before, we find that our coefficient of interest remains strong and is highly significant. These results remain robust and significant when we use our alternative measures of motivation for honesty in Panels B-D.

Overall, the resilience of these results corroborates our previous findings and demonstrates the importance of motivation for honesty in informal milk markets.

V. Conclusion

Understanding the determinants of large and persistent variation in cheating behavior in the face of asymmetric information and low monitoring is an important question in economics. Although laboratory and field evidence allude to the importance of intrinsic motivation for honesty, concrete evidence is lacking. A major challenge lies in developing measures that go beyond a simple classification of individuals into honest and dishonest types and capture the heterogeneity therein. In this paper, we develop novel measures that combine a standard die roll task with a Bluetooth enabled die which transmits actual outcomes of the die roll. This feature allows us to contrast self-reported with actual outcomes and develop refined measures of motivation for honesty that are less

[29]For example, we find that in the 710 rolls, in which a '5' was reported and the actual outcome was recorded, the share of over-reported rolls is 0.26. Thus, we add 0.26 to the number of over-reported rolls for each missed roll, in which a '5' was reported.

prone to measurement errors and capture not only the extensive but also the extensive margin. We then combine these measures with field outcomes on cheating to understand the importance of motivation for honesty for cheating behavior.

Our study takes place in a unique field setting comprising informal markets for buffalo milk in Delhi, India. These markets suffer from asymmetric information, as it is very difficult for customers to know ex-ante as well as to verify ex-post the true quality of milk. Moreover, there is no monitoring by third parties of these markets. These features provide milkmen strong incentives to dilute milk sold with water, leading primarily to nutritional losses.

Our study proceeds in three steps. We first invite milkmen to take part in the aforementioned behavioral experiment to measure their motivation for honesty. Our experimental findings reveal large heterogeneity in motivation for honesty: milkmen differ substantially in their willingness as well as the degree of dishonesty. A simple binary classification is unable to capture this heterogeneity and therefore suffers from sizable measurement errors. In the second step, we examine the importance of motivation for honesty for cheating in the field. We buy milk from the same milkmen and then gauge the amount of added water in milk using an ultrasonic milk analyzer. We find a strong and significant association between motivation for honesty and added water in milk, which holds to a powerful set of controls and a wide variety of robustness checks. In the final step, we show that varying degrees of dishonesty play a decisive role, such that when we contrast our estimates with those obtained from the binary classification, we find a large downward bias in the latter. These results support the complex and subtle nature of motivation for honesty, that it matters for explaining field outcomes on cheating, and that ignoring heterogeneity could undermine the explanatory power of motivation for honesty.

Our findings have important implications for public policy and future research. Heterogeneity in motivation for honesty calls for a differentiated policy response. Policies exclusively based on monitoring and punishment to counteract cheating might crowd-out intrinsic motivations (Bowles and Polonia-Reyes 2012) and thus could be combined with policies built on developing and fostering honest values. Integrating these findings into randomized control trials could be an important and interesting avenue to test for their relevance. Besides this, our study presents a challenge to the existing models of cheating that have thus far modeled the decision to cheat as a binary choice (Becker 1968, Kartik 2009, Olken and Pande 2012). These models predict that once the returns to cheating surpass a given individual-specific threshold, individuals would cheat maximally. We, however, do not find such a binary relationship between motivation for honesty and cheating. Our results rather suggest that there are also intermediate levels of dishonesty. This might reflect that internal costs of dishonesty are increasing in the magnitude of cheating. A better understanding of whats driving these differences in dishonesty is vital and an important area for future research.

References

Abeler, J., Becker, A., Falk, A., 2014. Representative Evidence on Lying Costs. *Journal of Public Economics* 113, 96-104.

Advanced Instruments. 1995. Added Water and the Freezing Point of Milk. Massachusetts: Advanced instruments, Inc.

Akerlof, G., 1970. The Market for Lemons: Qualitative Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84 (3), 488-500.

Andreoni, J. 1988. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37 (3), 291304.

Andreoni, J., Erard, B., Feinstein, J., 1998. Tax Compliance. *Journal of Economic Literature* 36 (2), 818-860.

Armantier, O., Boly, A., 2011. A Controlled Field Experiment on Corruption, *European Economic Review* 55 (8), 1072-82.

Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2013. What Drives Taxi Drivers? A Field Experiment on Fraud in a Market for Credence Goods. *Review of Economic Studies* 80 (3), 876-891.

Banerjee, A., Hanna, R., Mullainathan, S., 2012. Corruption. *Handbook of Organizational Economics*, ed. Gibbons, R. and J. Roberts, Vol 1, 1109-1147. New Jersey: Princeton University Press.

Bartling, B., Weber, R. A., Yao, L., 2015. Do Markets Erode Social Responsibility? *Quarterly Journal of Economics* 130 (1), 219-266.

Becker, G. S., 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2), 169 - 217.

Bowles, S., 1998. Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions. *Journal of Economic Literature* 36 (3): 75-111.

Bowles, S., and Polonia-Reyes, S. 2012. Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature* 50 (2): 368-425.

Cameron, C. A., Gelbach, J.B., Miller, D.B., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90 (3): 414-27.

CGWB, 2015. Contaminated Areas. New Delhi: Central Ground Water Board. `http://www.cgwb.gov.in/Documents/Contaminated%20Areas.pdf`, retrieved April 3rd, 2016.

Chaudhury, N., Hammer, J., Kremer, M, Muralidharan, M., Rogers, H., 2006. Missing in Action: Teacher and Health Worker Absence in Developing Countries, *Journal of Economic Perspectives, American Economic Association* 20 (1), 91-116.

Cohn, A., Fehr, E., Marchal, M., 2014. Business Culture and Dishonesty in the Banking Industry. *Nature*. doi:10.1038/nature13977.

Dai, Z., Galeotti, F., Villeval, M.C., 2016. Cheating in the Lab Predicts Fraud in the Field An Experiment in Public Transportations. Working Paper 1605, GATE, Lyon.

Di Tella, R., Perez-Truglia, R., Babino, A., Sigman, M., 2015. Conveniently Upset: Avoiding Altruism by Distorting Beliefs about Others' Altruism. *American Economic Review* 105 (11), 3416-42.

Duflo, E., Hanna, R., Ryan, S. P., 2012. Incentives Work: Getting Teachers to Come to School. *American Economic Review* 102 (4), 1241-78.

Dulleck, U., Kerschbamer, R., 2006. On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods. *Journal of Economic Literature* 44 (1), 5-42.

Dugar, S., Bhattacharya, H., 2016. Fishy Behavior: Evaluating Preferences for Honesty in the Marketplace. Mimeo. Calgary University.

Dulleck, U., Kerschbamer, R., Sutter, M., 2011. The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. *American Economic Review* 101 (2), 526-55.

FAO, 2013. Milk and Dairy Products in Human Nutrition.
`http://www.fao.org/docrep/018/i3396e/i3396e.pdf`.

Falk, A., Szech, N., (2013). Morals and markets. *Science* 340 (6133), 707-711.

Fischbacher, U., Foellmi-Heusi, F., 2013. Lies in Disguise. An Experimental Study on Cheating. *Journal of the European Economic Association* 11 (3), 525-547.

Fisman, R., Miguel, E., 2007. Corruption, Norms, and Legal Enforcement: Evidence From Diplomatic Parking Tickets. *Journal of Political Economy* 115 (6), 1020-1048.

Gaechter, S., Thoeni, C., 2005. Social Learning and Voluntary Cooperation Among Like-Minded People, *Journal of the European Economic Association* 3 (2-3), 303-314.

Gneezy, U., 2005. Deception: The Role of Consequences. *American Economic Review* 95 (1), 384-394.

Hanna, R., Wang, S.-Y., 2015. Dishonesty and Selection into Public Service. Working paper, Harvard University.

Hruschka, D., Efferson, C., Jiang, T., Falletta-Cowden, A., Sigurdsson, S., McNamara, R., Sands, M., Munira, S., Slingerland, E., Henrich, J., 2013. Impartial institutions, pathogen stress and the expanding social network.*Human Nature* 25 (4), 567-579.

Kartik, N., 2009. Strategic Communication With Lying Costs. *Review of Economic Studies* 76 (4), 1359-1395.

List, J. A., 2006. The Behavioralist Meets the Market: Measuring Social Preferences and

Reputation Effects in Actual Transactions. *Journal of Political Economy* 114 (1), 1-37.

List, J.A., 2009. The economics of open air markets. NBER Working Paper 15420.

Maggian, V. and M.C. Villeval, forthcoming. Social Preferences and Lying Aversion in Children. *Experimental Economics*. DOI 10.1007/s10683-015-9459-7.

Marai, I.F.M., Haeeb, A.A.M., 2010. Buffalo's Biological Functions as Affected by Heat Stress - A Review. *Livestock Science* 127 (2-3), 89-109.

Mazar, N., Amir, O., Ariely, D., 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research* 45 (6), 633-644.

NDDB, 2013-14. Annual Report. New Delhi: National Dairy Development Board. `http://www.nddb.org/sites/default/files/pdfs/nddb-annual-report-2013-2014.pdf`.

NSMA, 2011. Executive Summary on National Survey on Milk Adulteration. New Delhi: FSSAI.

Olken, B. A., 2007. Monitoring Corruption: Evidence From a Field Experiment in Indonesia. *Journal of Political Economy* 115, 200-249.

Olken, B. A., Pande, R., 2012. Corruption in Developing Countries. *Annual Review of Economics* 4 (1), 479-509.

Pruckner, G. J., Sausgruber, R., 2013. Honesty on the Streets: A Field Study on Newspaper Purchasing. *Journal of the European Economic Association* 11 (3), 661-679.

Rosenbaum, S. M., Billinger, S., Stieglitz, N., 2014. Let's Be Honest: A Review of Experimental Evidence of Honesty and Truth-Telling. *Journal of Economic Psychology* 45 (1), 181-196.

Shleifer, A., 2004. Does Competition Destroy Ethical Behavior? *American Economic Review* 94 (2), 414-418.

Sutter, M., 2009. Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams. *The Economic Journal* 119 (1), 47-60.

Yu. D., 2015. India's organized dairy sector anticipates rapid growth in next three years, http://www.dairyreporter.com/Markets/India-s-organized-dairy-sector-anticipates-rapid-growth-in-next-three-years, retrieved April 3rd, 2016.

Tables and Figures

Table 1: Summary Statistics

	Mean	Std. Dev.
	A: Field outcome on cheating	
Added water in milk	17.962	7.488
	B: Motivation for honesty	
Bluetooth-based measures:		
Number of over-reported rolls	3.625	6.604
Sum of added points	7.486	15.851
Share of over-reported rolls	10.710	19.543
Share of added points points	9.140	18.836
Conventional measures:		
Binary, 1-percent	0.125	0.333
Binary, 5-percent	0.222	0.419
Self-reported sum	147.472	21.504
	C: Controls	
Price	57.715	4.278
Socio-Demographics:		
Age	33.847	11.089
Education	10.319	3.439
English proficiency	2.167	0.856
Majority caste	0.486	0.503
Livestock-specific input factors:		
Buffalo herd size	18.778	35.619
Lactation period	0.792	0.409

Notes: Added water in milk is the percentage of water in 1 liter of buffalo milk purchased from the milkmen. Number of over-reported rolls is the number of rolls, in which the reported outcome exceeds the actual outcome. Sum of added points is the number of points added over all die rolls. Share of over-reported rolls is the ratio of over-reported rolls and the number of recorded rolls in which a milkman did not roll a '6' in percent. Share of added points is the ratio of the sum of added points and the maximum number of points, a milkman could have added given his actual outcomes in the recorded rolls in percent. Binary 1-percent is an indicator variable, which is 1 (dishonest) if a milkman reported 166 points or more, otherwise 0 (honest). Similarly, Binary 5-percent is an indicator variable, which is 1 (dishonest) if a milkman reported 158 points or more, otherwise 0 (honest). Self-reported sum is simply the sum of self-reported points over the rolls. Price is the amount paid for a liter of milk in Indian Rupees. Age is measured in years. Education is the years of schooling. English proficiency is the self-assessment of a milkman's knowledge of the English language on a 5-point scale, where 1 indicates poor knowledge and 5 indicates good knowledge. Majority caste is a dummy equal to 1, if a milkman belongs to the *gujjar* or *yadav* caste and 0 otherwise. Buffalo herd size is the number of adult buffaloes owned by a milkman. Lactation period is a dummy equal to 1, if any of their buffaloes were in the first three months of their lactation period in December 2014 and 0 otherwise. Data on lactation period was not available for two milkmen and was consequently imputed. The mean without the imputed values is 0.786 (s.d. 0.413). All our results hold, if we drop these observations.

Table 2: Heterogeneity in Patterns of Dishonesty

	Number of over-reported rolls (1)	(2)	(3)	Sum of added points (4)	Share of over-reported rolls (5)	Share of added points (6)
Strong Dishonesty	17.444***	17.352***	17.115***	36.986***	51.099***	47.604***
	(2.508)	(2.575)	(2.635)	(8.181)	(7.193)	(8.384)
Weak Dishonesty	3.852***	3.923***	3.824***	6.575***	11.388***	7.750***
	(0.771)	(0.837)	(0.934)	(2.074)	(2.991)	(2.380)
Age		0.023	0.015	0.041	0.041	0.020
		(0.035)	(0.033)	(0.096)	(0.102)	(0.099)
Education		0.030	−0.035	−0.325	−0.021	−0.107
		(0.199)	(0.181)	(0.569)	(0.487)	(0.516)
English proficiency		−0.017	0.050	1.505	−0.699	0.122
		(1.086)	(1.104)	(3.794)	(3.226)	(3.606)
Majority caste		0.472	−0.115	−0.400	−0.570	−0.469
		(0.835)	(0.866)	(2.460)	(2.528)	(2.747)
Buffalo herd size		−0.006	−0.005	−0.012	−0.007	−0.004
		(0.011)	(0.011)	(0.037)	(0.032)	(0.037)
Lactation period		0.790	0.861	3.299	2.253	3.504
		(0.804)	(0.872)	(2.694)	(2.324)	(2.660)
Constant	0.000	−1.817	−1.393	−6.069	−1.524	−3.740
	(.)	(2.897)	(3.181)	(8.200)	(9.069)	(8.104)
Fixed Effects	No	No	Yes	Yes	Yes	Yes
Observations	72	72	72	72	72	72
R^2	0.708	0.714	0.726	0.640	0.724	0.715

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is number of over-reported rolls in columns (1) to (3), sum of added points in column (4), share of over-reported rolls in column (5), and share of added points in column (6). Controls are as defined in the footnote of Table 1. Strong Dishonesty is a binary variable equal to 1 for milkmen whose self-reported sum is 166 points or more, and 0 otherwise. Weak Dishonesty is a binary variable equal to 1 for milkmen whose number of over-reported rolls is larger than 0 but whose self-reported sum is less than 166 points. Fixed effects include dummy variables for the six dairies.
*** Significant at the 1 percent level

Table 3: Motivation for Honesty and Cheating in Milk Markets

	Added water in percent				Protein	SNF
	No Control	Price	Controls	F.E.	F.E.	F.E.
	(1)	(2)	(3)	(4)	(5)	(6)
Dishonesty	0.291*	0.286**	0.315***	0.442***	−0.062***	−0.029***
	(0.158)	(0.133)	(0.108)	(0.117)	(0.017)	(0.008)
Price		−0.625***	−0.568***	−0.471**	0.064**	0.031**
		(0.176)	(0.203)	(0.223)	(0.032)	(0.015)
Age			−0.004	−0.006	0.002	0.000
			(0.070)	(0.062)	(0.008)	(0.004)
Education			0.256	0.479	−0.063	−0.030
			(0.295)	(0.331)	(0.045)	(0.021)
English proficiency			−1.713	−2.226*	0.289*	0.132*
			(1.338)	(1.204)	(0.169)	(0.078)
Majority caste			2.531	3.038	−0.438	−0.202
			(1.631)	(2.256)	(0.313)	(0.146)
Buffalo herd size			0.020	0.022	−0.003	−0.001
			(0.014)	(0.015)	(0.002)	(0.001)
Lactation period			−5.412**	−5.564***	0.723**	0.341***
			(2.159)	(1.937)	(0.275)	(0.128)
Constant	16.905***	53.025***	53.498***	48.654***	3.035	1.271
	(0.944)	(10.120)	(12.062)	(13.755)	(1.924)	(0.885)
Fixed Effects	No	No	No	Yes	Yes	Yes
Observations	72	72	72	72	72	72
R^2	0.066	0.194	0.334	0.540	0.527	0.529

Notes: OLS regression with robust standard errors in parentheses. The dependent variable in columns (1)-(4) is added water in milk in percent. In column (5), the dependent variable is protein in percent. In column (6), the dependent variable is solids-not-fat (SNF) in percent. Controls are as defined in the footnote of Table 1. Motivation for honesty is measured via dishonesty, which is the number of over-reported rolls. Fixed effects include dummy variables for the time of day milk was purchased, dairies, and assistants.

* Significant at the 10 percent level

** Significant at the 5 percent level

*** Significant at the 1 percent level

Table 4: Degrees of Dishonesty and Measurement Error

	Added Water in Milk in Percent						
	Bluetooth Measures			Conventional Measures			
	Intensive margin	Extensive Margin 1 percent	Extensive Margin 5 percent	Binary 1 percent	Binary 5 percent	Self-reported Sum	Self-reported Sum + Actual
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Strong dishonesty	0.480***	9.571***	4.869**	8.007***	3.587*		
	(0.124)	(2.537)	(2.163)	(2.353)	(1.937)		
Weak dishonesty	0.260*	3.139**	2.836*				
	(0.146)	(1.419)	(1.614)				
Self-reported sum						0.109***	0.143***
						(0.039)	(0.039)
Actual outcome							−4.160
							(3.189)
Price	−0.455**	−0.546**	−0.567**	−0.489**	−0.514*	−0.541**	−0.496*
	(0.220)	(0.219)	(0.245)	(0.239)	(0.262)	(0.245)	(0.260)
Age	−0.003	0.021	−0.035	0.000	−0.046	−0.047	−0.031
	(0.063)	(0.067)	(0.064)	(0.063)	(0.065)	(0.064)	(0.066)
Education	0.452	0.552	0.522	0.404	0.399	0.357	0.437
	(0.335)	(0.342)	(0.343)	(0.328)	(0.334)	(0.342)	(0.338)
English proficiency	−2.291*	−2.492**	−2.285*	−2.184*	−2.001	−2.263*	−2.438**
	(1.224)	(1.160)	(1.248)	(1.161)	(1.263)	(1.148)	(1.156)
Majority caste	2.984	2.998	2.569	2.901	2.562	2.045	2.630
	(2.333)	(2.186)	(2.434)	(2.300)	(2.485)	(2.466)	(2.499)
Buffalo herd size	0.024	0.022	0.011	0.026*	0.015	0.019	0.020
	(0.015)	(0.014)	(0.015)	(0.014)	(0.016)	(0.015)	(0.015)
Lactation period	−5.726***	−5.601***	−5.034**	−5.477***	−4.993**	−5.639***	−5.753***
	(1.912)	(2.008)	(2.154)	(1.919)	(2.078)	(2.038)	(2.012)
Constant	48.489***	51.192***	55.143***	50.311***	53.904***	42.201***	48.198***
	(13.712)	(13.123)	(14.579)	(14.453)	(15.714)	(14.517)	(14.245)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	72	72	72	72	72	72	72
R^2	0.546	0.558	0.483	0.524	0.459	0.497	0.513

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. In column (1), strong dishonesty is the number of over-reported rolls for the nine milkmen whose self-reported sum is 166 points or more, and 0 otherwise. In columns (2) and (4), strong dishonesty is a binary variable, which is equal to 1 for milkmen whose self-reported sum is 166 points or more, and 0 otherwise. In columns (3) and (5) strong dishonesty is a binary variable, which is equal to 1 for milkmen whose self-reported sum is 158 points or more, and 0 otherwise. In column (1), weak dishonesty is the number of over-reported rolls for milkmen whose number of self-reported points is less than 166, and 0 otherwise. In column (2), weak dishonesty is a binary variable, which is equal to 1 for milkmen whose number of over-reported rolls is larger than 0 and whose number of self-reported points is less than 166, and 0 otherwise. In column (3), weak dishonesty is a binary variable, which is equal to 1 for milkmen whose number of over-reported rolls is larger than 0 and whose number of self-reported points is less than 158, and 0 otherwise. Self-reported sum is the number of reported points over all 40 die rolls. Actual outcome is the average number transmitted by the bluetooth die over all recorded rolls. Controls are as defined in the footnote of Table 1. Fixed effects include dummy variables for the time of day milk was purchased, the six dairies, and assistants.

* Significant at the 10 percent level
** Significant at the 5 percent level
*** Significant at the 1 percent level

Table 5: Alternative Measures of Motivation for Honesty

	Added Water in Milk in Percent		
	Sum of added points (1)	Share of over-reported rolls (2)	Share of added points (3)
Dishonesty	0.159***	0.146***	0.148***
	(0.044)	(0.039)	(0.036)
Price	−0.473*	−0.476**	−0.463*
	(0.244)	(0.230)	(0.235)
Age	−0.019	−0.006	−0.008
	(0.062)	(0.062)	(0.062)
Education	0.466	0.462	0.437
	(0.335)	(0.329)	(0.333)
English proficiency	−2.297*	−2.092*	−2.121*
	(1.221)	(1.193)	(1.203)
Majority caste	2.932	3.057	3.023
	(2.348)	(2.251)	(2.310)
Buffalo herd size	0.020	0.021	0.020
	(0.015)	(0.015)	(0.015)
Lactation period	−5.386**	−5.482***	−5.530***
	(2.043)	(1.953)	(1.980)
Constant	50.236***	48.685***	49.019***
	(14.709)	(14.083)	(14.263)
Std. Coefficient	2.592	2.856	2.787
Fixed Effects	Yes	Yes	Yes
Observations	72	72	72
R^2	0.506	0.537	0.527

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. Controls are as defined in the footnote of Table 1. Fixed effects include dummy variables for the time of day milk was purchased, dairies, and assistants.
* Significant at the 10 percent level
** Significant at the 5 percent level
*** Significant at the 1 percent level

Table 6: Additional Controls

	Added Water in Milk in Percent	
	Additional Variable	Number of over-reported rolls
	(1)	(2)
Buffaloes from Haryana	−0.404**	0.428***
	(0.184)	(0.100)
Buffalo feed costs	0.322	0.462***
	(0.235)	(0.112)
Duration of stay	−0.106	0.441***
	(0.092)	(0.119)
State of origin	−1.271	0.435***
	(1.823)	(0.116)
Outside option	−0.872	0.435***
	(1.613)	(0.124)
Religiosity	−0.019	0.451***
	(0.020)	(0.118)
Family size	−0.136	0.429***
	(0.303)	(0.123)
Dishonesty neighbor	0.071	0.455***
	(0.182)	(0.120)

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. Controls include age, education, English proficiency, majority caste, buffalo herd size, and lactation period, and are defined in the footnote of Table 1. Additional controls are defined in section IV. Buffalo feed costs are measured in 1,000 INR. Fixed effects include dummy variables for the time of day milk was purchased, dairies, and assistants. The number of observations is 72, except in row 1 where it is 64, row 2 where it is 63, and row 3 where it is 71.

** Significant at the 5 percent level

*** Significant at the 1 percent level

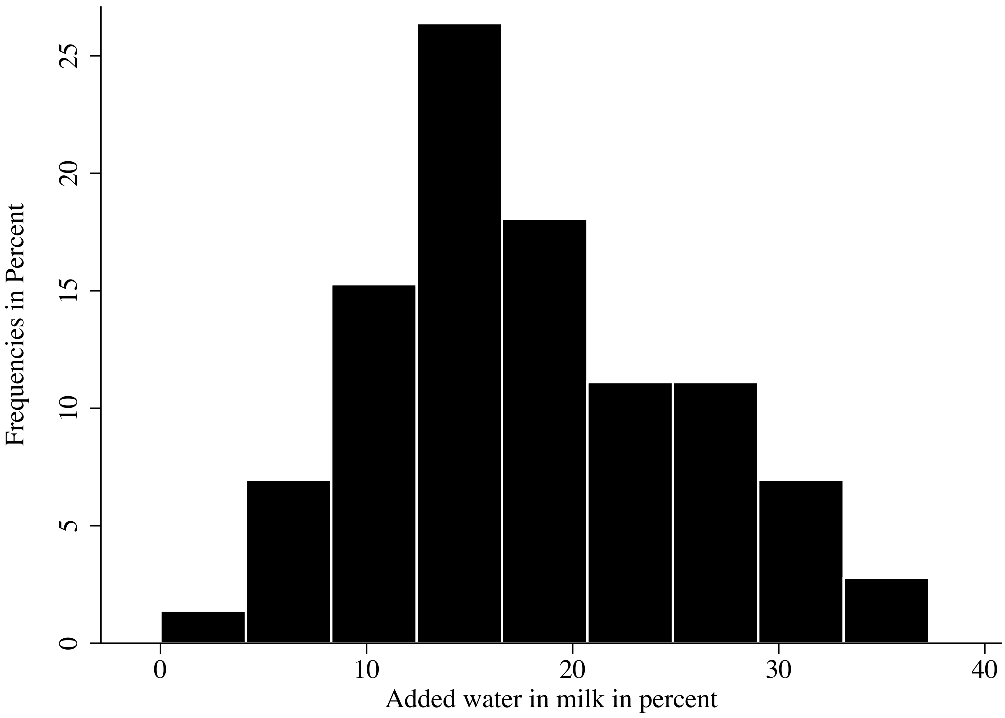


Figure 1: Distribution of Added Water in Milk

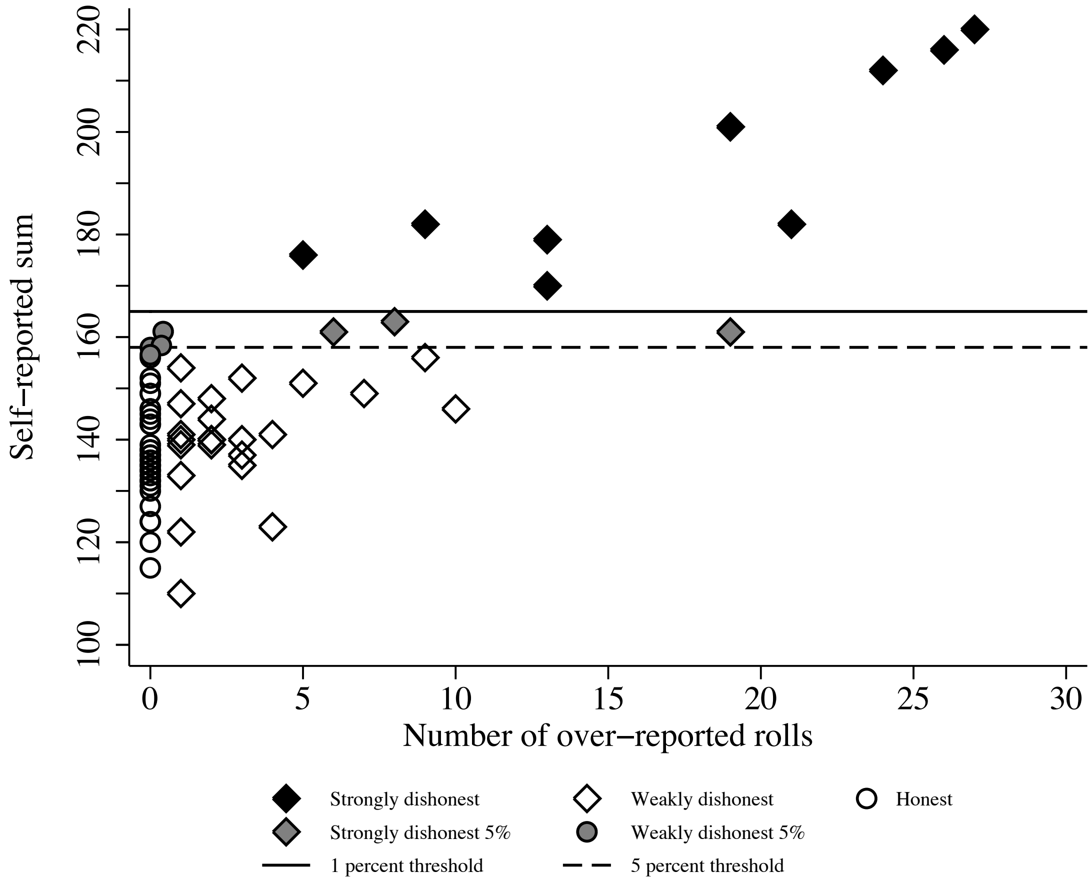


Figure 2: Heterogeneity and Measurement Error in Motivation for Honesty

Notes: The reported sum over all 40 die rolls is depicted on the y-axis, the number of over-reported rolls on the x-axis. The solid line shows the critical threshold for binary classification at the 1 percent level of significance (166 points), the dashed for the 5 percent level (158 points). Circles represent those milkmen who did not over-report in the experiment, squares those milkmen who did. Milkmen who are classified as "dishonest" at the 1-percent level of significance are marked black, those who are additionally classified as "dishonest" at the 5 percent level are marked grey.

ONLINE APPENDIX:
Measuring Motivation for Honesty and Explaining Cheating in Informal Milk Markets in India

Markus Kröll and Devesh Rustagi

Appendix A. Additional results

I. Procedures for Collecting the Main Outcome

In this section, we provide a description of the procedures for collecting the main outcome on cheating, which proceeds in two steps.

Pilot study

In the first step, we conducted a pilot study to identify different kinds of adulterants that milkmen may add to milk, as well as to validate the measure of added water in milk provided by the ultrasonic milk analyzer. In a nation-wide study conducted by FSSAI, many different adulterants were found in milk samples, but in Delhi it was primarily water (NSMA 2011). Nonetheless, to avoid underestimating the extent of cheating by focussing only on added water, we collected milk samples from 105 milkmen from dairies in our study and tested these samples for a broader set of adulterants listed in the FSSAI study. We split each sample into two parts. One part was sent to a professional food-testing laboratory in Delhi (Sima Lab Pvt Ltd.) to test for the presence of water, starch, urea, detergent, skimmed milk powder, and glucose. The second part was tested only for added water using the milk analyzer because the machine is unable to detect other adulterants. Mirroring the results of the FSSAI-study, these analyses revealed that water is the main adulterant of milk in Delhi. Moreover, the correlation between estimates on added water by the laboratory and the machine is very strong ($r = 0.93$). As a result, we focused on added water in buffalo milk in percent measured using the machine as our field outcome on cheating.

We rely on the machine measure because it allows for a more flexible, cheaper, and precise analysis of added water. While the laboratory imposed a limit of 20 samples per week, charged INR 1,250 per sample, used a lactometer, provided mostly qualitative results, and took a week to deliver the results, the machine took only two minutes per sample to give the results.

Sampling procedures for the assessment of main outcome

Milk samples were collected early in the morning (around 7 am) and in the afternoon (around 4 pm), shortly after the buffaloes are milked, in the third week of December. We hired five assistants unknown to the milkmen to execute this task.

In each dairy, every assistant purchased a liter of milk from five to eight milkmen spread out over several shifts. The set of milkmen for each assistant was assigned such that further contact with a given milkman was avoided after milk was bought from him. The purchased milk was then brought to a car outside the dairy and transferred into a clearly labeled plastic bottle, which contained a specific identification number for every milkmen. These bottles were then stored in an ice box to prevent spoilage.

A major concern in collecting these milk samples is locating the farms of the milkmen who took part in our experiment, because most of the six dairies resemble informal settlements. Since most farms do not have plot numbers written outside their house mismatches can easily occur. In order to avoid this problem, we prepared detailed maps of each dairy so that assistants could accurately locate the milkmen we wanted to target.

We prepared these maps (see Fig. A.1) through guided walks and photographs while conducting the second household survey. We marked every target milkmen on the map and gave notice about the location (e.g. color of the house, nearby shops, signs, and pole numbers, etc.). For particularly difficult matches, assistants were requested to take pictures of the dairy farm with their mobile phones, which were subsequently verified using pictures independently obtained by us.

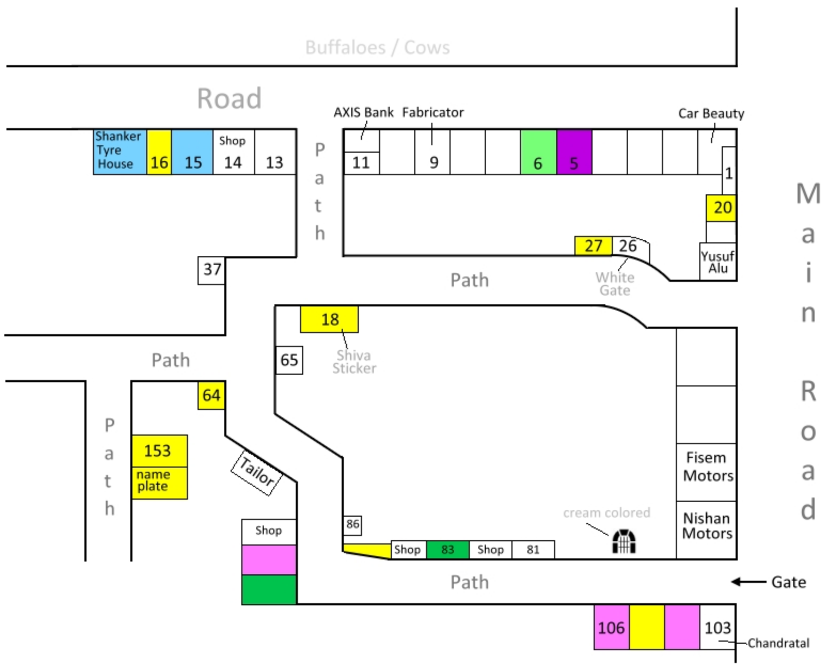


Figure A.1: Sample Map

II. Sample Construction

Selection into the experiment

To control for the potential selection of milkmen into our experiment, we collected milk samples from 63 milkmen who did not take part in our experiment. These milkmen reside in the same dairies and also operate small farms themselves just like the 72 milkmen who took part in the experiment. The distribution of added water in milk does not differ between these two groups of milkmen (*Kolmogorov-Smirnov*: $p = 0.58$). Test statistics based on a comparison of means (Non-participants: 19.05, participants: 17.96, p-value = 0.42) and the median (Non-participants: 15.92, participants: 16.72, p-value= 0.64) yield similar results.

Sample attrition

Despite our best efforts in collecting a milk sample from every participating milkman in our experiment, we could not obtain the field outcome for nine milkmen. In addition, we dropped three milkmen from our analysis for whom the actual outcomes in the experiment are missing due to outages in the Bluetooth-connection. Table A1 demonstrates that these 12 milkmen do not differ significantly from the 72 milkmen in our main sample in key socio-demographic characteristics.

Table A.1: Testing for Sample Selection

	Participants in the sample (1)	Participants not in the Sample (2)	Difference (3)
Self-reported sum	147.472	145.25	2.222
	(2.534)	(6.314)	(6.721)
Age	33.847	32.583	1.264
	(1.307)	(2.638)	(3.383)
Education	10.319	10.25	0.069
	(0.405)	(0.986)	(1.071)
English proficiency	2.167	2.333	−0.167
	(0.101)	(0.225)	(0.264)
Majority caste	0.514	0.417	0.097
	(0.059)	(0.142)	(0.756)
Buffalo herd size	18.778	9.25	9.528
	(4.198)	(3.572)	(10.431)
Lactation period	0.208	0.444	−0.236
	(0.048)	(0.166)	(0.203)

Notes: Comparison of the characteristics of the 72 milkmen in our final sample and the 12 milkmen we dropped. Variables are as defined in the footnote of Table 1 of the main paper. Column (3) reports the difference between the two groups and its standard errors in parentheses. Differences are tested for significance using a two-sided t-test. Differences in the variables majority caste and lactation period are tested for significance using a two-sided Fisher-Exact test.

III. Milk-Testing Tournament

We examine the reliability of informal tests, which can be easily implemented at home, to detect added water in milk. Our rationale is that if such household tests are able to rate the quality of milk, reputation and product differentiation could potentially mitigate cheating due to asymmetric information. To this end, we conducted an incentivized milk-testing tournament to assess the scope of such household remedies.

For this experiment, we bought a liter of pure buffalo milk from milkmen who took part in our study. We then created five different milk samples by adding varying levels of water to the pure milk, ranging from 0 to 400 milliliters, in the units of 100 milliliters. We then presented these five milk samples to milkmen and asked them to rate each sample on the amount of added water. 74 milkmen from four of the six dairies in our sample took part in this experiment. The three milkmen in each dairy whose combined assessment was closest to the actual water levels were paid INR 800, INR 500, and INR 300 respectively, which is a fairly high stake. We chose to conduct the experiment with milkmen because they produce, dilute, and consume milk every day, and thus are experts in judging milk quality.

Figure A1 reports the results and shows that verification of milk quality using simple test procedures, such as tasting milk or testing its viscosity, is extremely difficult. It is evident that milkmen are unable to distinguish milk samples even at the widest margin. Irrespective of the actual water in milk, each of the five samples was considered by some milkmen to be pure, while others predicted the same sample to contain water in excess of 50 percent. As the inter-quartile range reveals, this wide range is not because of some outliers. Notably, the median predicted water level for pure milk exceeds the corresponding values for the most diluted sample. Moreover, regressing the predicted water level on the actual amount of added water, we find a negative albeit insignificant coefficient (-0.022, p-value = 0.585). These results clearly demonstrate that the lack of ex-post verifiability barring expensive and for the public inaccessible methods (such as professional testing laboratories) severely hampers the scope for reputation or product differentiation to mitigate cheating in these markets.

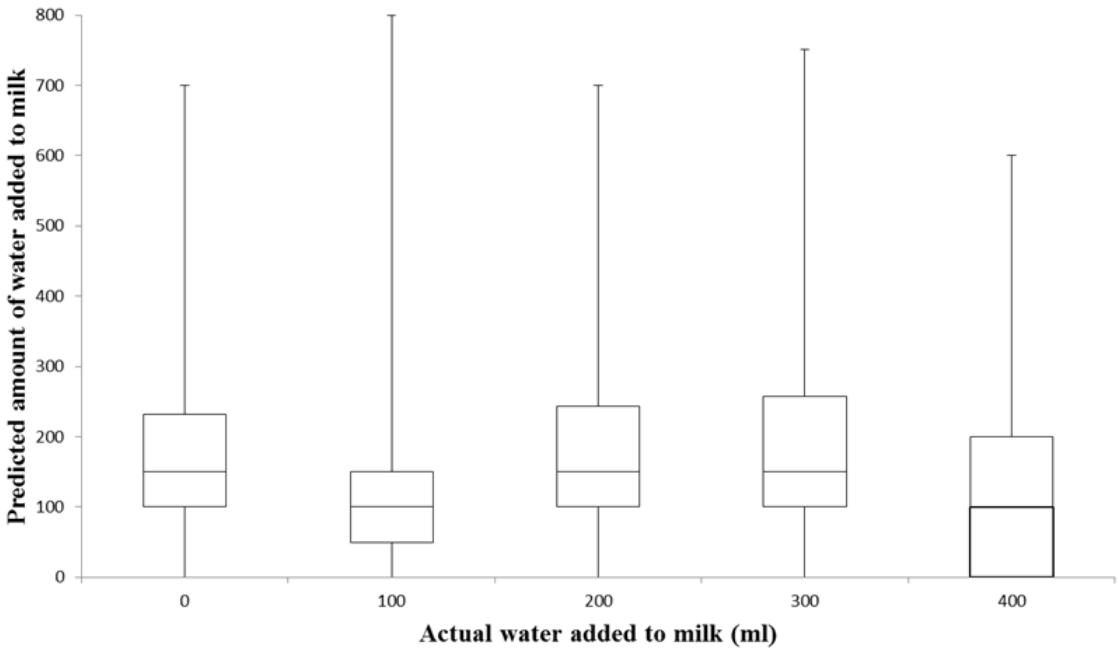


Figure A.2: Milk-testing tournament

IV. Control Experiment

To control for a potential effect of the Bluetooth die itself on cheating behavior, we conducted a control experiment with 105 participants in which one half played the game with the Bluetooth die and the other half was given a regular die. Comparing the reported sum of points for the two groups reveals that the group which used the Bluetooth die does not differ from the group which played the game with a regular die. Testing for differences in the distribution across the two groups with a Kolmogorov-Smirnov test, we find no significant difference (p-value= 0.698). Similarly, the mean across the group which used the regular die (158.88 points) is not significantly different from the mean across the group that used the Bluetooth die (153.49); the difference being 5.39 points, (p-value= 0.343).

V. Motivation for Honesty and Control Variables

Table A2 reports the pairwise correlation using the Pearson product-moment correlation coefficient between our continuous measure of motivation for honesty and the covariates employed in our regression specification. p-values are reported in parentheses. With the exception of family size (at the 10-percent level of significance), no control variable is statistically significantly correlated with the number of over-reported rolls.

Table A.2: Correlation of Motivation for Honesty and Other Covariates

	Controls						
	Price	Age	Education	English proficiency	Majority caste	Buffalo herd size	Lactation period
Number of over-reported rolls	−0.014	−0.084	0.048	0.098	0.149	−0.029	0.117
	(0.908)	(0.486)	(0.688)	(0.411)	(0.212)	(0.809)	(0.329)

Additional controls							
Buffalo from Haryana	Buffalo feed costs	Duration of stay	State of origin	Outside option	Religiosity	Family size	Dishonesty neighbor
0.116	0.013	0.015	−0.142	−0.039	0.037	−0.211*	−0.190
(0.360)	(0.919)	(0.903)	(0.235)	(0.747)	(0.759)	(0.076)	(0.110)

Notes: Pairwise correlation between the number of over-reported rolls and all covariates. p-values in parentheses. Controls are as defined in the footnote of Table 1, additional controls are as defined in section VI of Appendix A.
* Significant at the 10 percent level

VI. Additional Regression Tables

In this section, we report further robustness checks of our main results in section III.

We first show in Table A.3 that the inclusion of the additional control variables also does not affect the association between our alternative measures of motivation for honesty and cheating in the field. The alternative measures used are sum of added points in column 2, share of over-reported rolls in column 4, and share of added points in column 6. The coefficient on the added variable is listed in the column preceding the alternative measure.

Table A.3: Robustness Check: Additional Controls and Alternative Measures

	Added Water in Milk in Percent					
	Additional Variable	Sum of added points	Additional Variable	Share of over-reported rolls	Additional Variable	Share of added points
	(1)	(2)	(3)	(4)	(5)	(6)
Buffaloes from Haryana	−0.398**	0.157***	−0.396**	0.141***	−0.392**	0.142***
	(0.186)	(0.037)	(0.186)	(0.034)	(0.185)	(0.032)
Buffalo feed costs	0.373	0.173***	0.336	0.156***	0.374	0.159***
	(0.247)	(0.046)	(0.237)	(0.037)	(0.245)	(0.038)
Duration of stay	−0.101	0.158***	−0.103	0.145***	−0.096	0.146***
	(0.095)	(0.046)	(0.093)	(0.040)	(0.095)	(0.038)
State of origin	−1.776	0.158***	−1.263	0.144***	−1.553	0.146***
	(1.884)	(0.043)	(1.830)	(0.039)	(1.861)	(0.036)
Outside option	−0.717	0.155***	−0.917	0.144***	−0.874	0.145***
	(1.672)	(0.048)	(1.607)	(0.041)	(1.634)	(0.040)
Religiosity	−0.015	0.161***	−0.021	0.150***	−0.016	0.150***
	(0.020)	(0.045)	(0.020)	(0.039)	(0.019)	(0.037)
Family size	−0.159	0.152***	−0.158	0.141***	−0.143	0.143***
	(0.307)	(0.046)	(0.302)	(0.041)	(0.307)	(0.039)
Dishonesty neighbor	0.061	0.164***	0.074	0.151***	0.070	0.152***
	(0.183)	(0.046)	(0.183)	(0.040)	(0.183)	(0.038)

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. Controls include age, education, English proficiency, majority caste, buffalo herd size, and lactation period, and are defined in the footnote of Table 1. Additional controls are defined in section IV. Buffalo feed costs are measured in 1,000 INR. Fixed effects include dummy variables for the time of day milk was purchased, dairies, and assistants. The number of observations is 72, except in row 1 where it is 64, row 2 where it is 63, and row 3 where it is 71.
** Significant at the 5 percent level
*** Significant at the 1 percent level

We next establish that our results are not driven by any particular dairy or due to influential observations. Table A.4 shows the corresponding changes in motivation for honesty, when we drop one dairy at a time or six influential observations across dairies respectively. Influential observations are determined using DFITS. An observation is classified as influential, if the difference in fitted values with and without the observation is larger than $2 \cdot \sqrt{\frac{k}{N}}$, where k is the number of parameters and N is the sample size.

Table A.4: Robustness Check: Influential Observations

	Added Water in Milk in Percent			
	Number of over-reported rolls	Sum of added point	Share of over-reported rolls	Share of added points
	(1)	2	(3)	(4)
Baseline	0.442***	0.159***	0.146***	0.148***
	(0.117)	(0.044)	(0.039)	(0.036)
Dairy 1	0.467***	0.153***	0.162***	0.148***
	(0.113)	(0.042)	(0.038)	(0.036)
Dairy 2	0.450***	0.162***	0.149***	0.151***
	(0.126)	(0.046)	(0.043)	(0.039)
Dairy 3	0.437**	0.151**	0.143**	0.147**
	(0.168)	(0.064)	(0.055)	(0.058)
Dairy 4	0.520***	0.197***	0.169***	0.176***
	(0.134)	(0.052)	(0.044)	(0.042)
Dairy 5	0.307**	0.126*	0.100**	0.108**
	(0.141)	(0.068)	(0.044)	(0.049)
Dairy 6	0.444***	0.159***	0.146***	0.148***
	(0.116)	(0.042)	(0.039)	(0.036)
Influential	0.329***	0.126***	0.110***	0.113***
Observations	(0.096)	(0.042)	(0.032)	(0.032)

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. Controls include age, education, English proficiency, majority caste, buffalo herd size, and lactation period as defined in the footnote of Table 1. Fixed effects include dummy variables for the time of day milk was purchased, the six dairies, and assistants.
* Significant at the 10 percent level
** Significant at the 5 percent level
*** Significant at the 1 percent level

Table A.5 reports a battery of robustness checks concerning Bluetooth misses. In column 1, each observation of the number of over-reported rolls is weighted by the number of recorded rolls. Column 2 includes the number of missed rolls, i.e. the number of rolls in which the bluetooth die did not transmit the outcome and an outcome larger than '1' was reported, as an additional co-variate. All milkmen who had more than 10 missed rolls are dropped in column 3. For column 4, all missed rolls are added to the number of over-reported rolls. In column 5, the number of over-reported rolls is weighted linearly by the number of missed rolls. For column 6, we compute the share of over-reported rolls for each outcome of the die roll based on the observed behavior in the recorded rolls in our entire sample, and add the respective share to the number of over-reported rolls for the outcome of each missed roll. Panel A reports the results for the number of over-reported rolls, Panel B reports the results for the sum of added points, Panel C for the share of over-reported rolls, and Panel D for the share of added points.

Table A.5: Robustness Check: Missed Rolls

	Added Water in Milk in Percent					
	Weighted regression	Missed rolls	Dropped if missed rolls ≥ 10	Maximum penalty	Linearly weighted	Observed behavior
	Panel A: Number of over-reported rolls					
Number of over-reported rolls	0.467*** (0.127)	0.457*** (0.127)	0.446*** (0.127)	0.280*** (0.087)	0.365*** (0.097)	0.431*** (0.112)
Missed rolls		−0.114 (0.201)				
	Panel B: Sum of added points					
Sum of added points	0.159*** (0.050)	0.162*** (0.047)	0.146*** (0.050)	0.091*** (0.032)	0.137*** (0.036)	0.156*** (0.042)
Missed rolls		−0.057 (0.207)				
	Panel C: Share of over-reported rolls					
Share of over-reported rolls	0.159*** (0.042)	0.155*** (0.043)	0.154*** (0.043)	0.073** (0.029)	0.117*** (0.032)	0.140*** (0.037)
Missed rolls		−0.164 (0.202)				
	Panel D: Share of added points					
Share of added points	0.154*** (0.042)	0.155*** (0.042)	0.142*** (0.041)	0.052* (0.026)	0.122*** (0.029)	0.142*** (0.034)
Missed rolls		−0.139 (0.203)				
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	72	72	63	72	72	72

Notes: OLS regression with robust standard errors in parentheses. The dependent variable is added water in milk in percent. In column 1, each observation of the number of over-reported rolls is weighted by the number of recorded rolls. Column 2 includes the number of missed rolls, i.e. the number of rolls in which the bluetooth die did not transmit the outcome and an outcome larger than '1' was reported, as an additional co-variate. All milkmen who had more than 10 missed rolls are dropped in column 3. For column 4, all missed rolls are added to the number of over-reported rolls. In column 5, the number of over-reported rolls is weighted linearly by the number of missed rolls. For column 6, we compute the share of over-reported rolls for each outcome of the die roll based on the observed behavior in the recorded rolls in our entire sample, and add the respective share to the number of over-reported rolls for the outcome of each missed roll. Controls include age, education, English proficiency, majority caste, buffalo herd size, and lactation period as defined in the footnote of Table 1. Fixed effects include dummy variables for the time of day milk was purchased, the six dairies, and assistants.
* Significant at the 10 percent level
** Significant at the 5 percent level
*** Significant at the 1 percent level

Appendix B. Experimental Instructions and Procedures

I. General Instructions (translated from Hindi)

Greetings and welcome to all of you. My name is Devesh Rustagi and his name is Markus Kröll. We are working at a university in Germany. We are here for research on the livelihood of milkmen. We hope that you will help us with our study. Please switch off your cell phones now. We thank you a lot for your support.

1. In this research, we would like to play a few games with you. In these games, you can earn some money. How much you can earn depends on how you play the game.

2. In the games, your identity will be kept anonymous. I am interested only in the decisions made by you in these games and not your identity. This is the reason that we removed your name from your personal invitation card. We will identify your decision in the game with a sticker like this (*show sticker*). You will draw a sticker like this from a lottery and we will stick it to your personal invitation card. Please do not lose the invitation card.

3. We will play two different games with you. You can earn money in both the games, which we will pay you immediately after the games are over.

4. We will give you separate instructions on how to play each game. Before we play the first game, we will give you the instructions on how to play the first game. Likewise, when we play the second game, we will give you the instructions for the second game. It is very important that you listen to these instructions carefully. In case you do not understand the game, please do not hesitate to ask us. We will be happy to assist you.

5. Before the start of the actual game, we will ask some questions to verify that you have understood the game. Therefore, it is important that you pay attention to our explanations and instructions.

6. Please do not discuss the games with the other players.

7. Do you have any questions as of now? If not, then we will begin with the instructions for the first game.

II. Instructions for the Die Game (translated from Hindi)

Instructions die game

1. You play this game on your own.
2. We will give you a die like this (*show the die*) and a sheet of paper like this (*show the sheet*).
3. All you have to do is to roll the die and report the number on the sheet of paper which we gave you.
4. To record the number, please cross INR 2 coins in the appropriate row. Each row has 6 coins, one coin for each point on the die (*show it on the poster*).
5. You get 2 INR for each coin you cross. Let?s take some examples:
   Example 1: if you cross 2 coins, then we will pay you INR 4 (*cross two coins on the poster*);
   Example 2: if you cross 5 coins, then we will pay you INR 10 (*cross five coins on the poster*).
6. You will have to repeat this task 40 times.
7. Your final earnings for this game will be the sum of earnings in each of the 40 rounds. We will sum the total earning over all rounds for you.
8. This means, the minimum you can earn is 80 INR and the maximum is 480 INR.
9. You will play this game in a private cabin (*show the cabin*). Once you are done playing this game, please give the sheet to us.
10. Please leave the room. We will call you one by one.

Control questions ***(Individually)***

Do you have any further questions? If no, we will ask you a few control questions.

1. How many times do you roll the die?
2. How much money do you earn by crossing a coin?
3. How is your income calculated?

Procedure ***(Individually)***

1. Please roll the die like this on the table (*Demonstrate proper die roll)*
2. If the die drops off the table, pelase do not record the outcome and repeat the die roll. Please make sure that the die does not drop-
3. After the game is over, please give us the sheet.

III. Experimental Procedures

In the following we briefly outline the procedural details of our experiment. The experiment was conducted within the premises of each dairy neighborhood a month before we collected the milk samples for the final field outcome. The experiment was scheduled such that participation did not overlap with the daily business of milkmen. We personally notified selected milkmen a few days before the experiment with the help of a community mobilizer from the respective dairy neighborhood. All selected milkmen were given a personalized invitation card containing their individual plot number in the dairy, which served as an admission to the experiment (Figure B1, top card). In addition, these cards enabled us to match experimental and field outcome. Each card had a unique ID number written on its back using a UV-readable pen. Thus, these IDs were invisible to the milkmen (Figure B1, middle card) and was only readable using UV-light (Figure B1, bottom card). We verified in the post-game interviews that milkmen did not exchange these invitation cards.

On the day of the experiment, we first carefully explained the purpose and procedure of the experiment at the group level. Each milkman then replaced his individual plot ID number with an identity card of his own choice bearing the names of European states (see Figure B1, top card). We then gave detailed instructions and examples at the group level for our die-game that were tested and polished in four pilot studies. Following these group-level instructions, every milkman was individually led into a room in which the experiment took place.

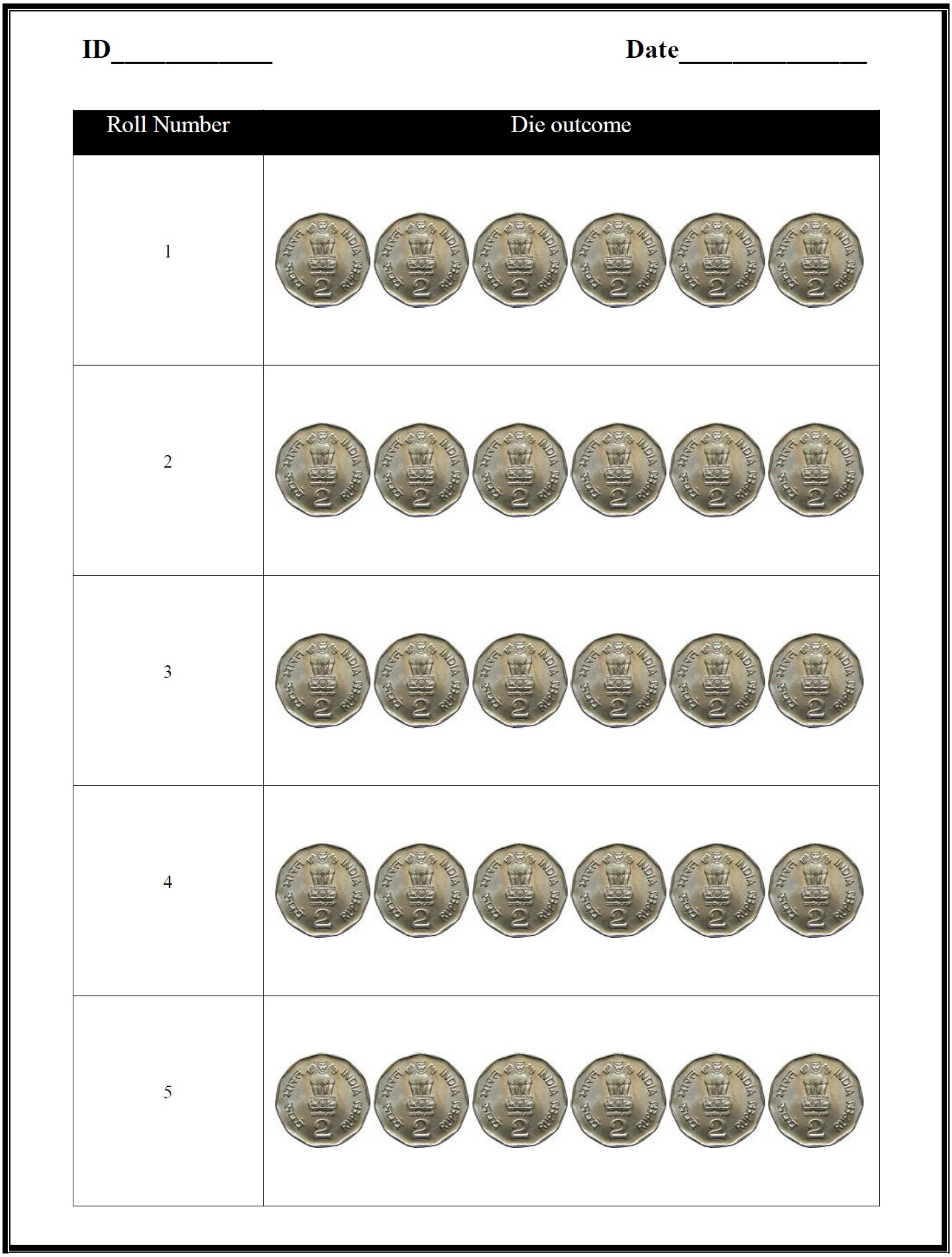
Figure B1: Invitation card

For the actual game we undertook great efforts to create the impression of full privacy: every participant was individually led into a room where they carried out the task on their own. Participants were not informed about the Bluetooth die (Figure B2) and thus operated under the belief of no scrutiny. Before milkmen took part in the actual experiment, they had to answer three control questions correctly and were once again shown how to roll the die. This individual demonstration was implemented in order to minimize deliberate manipulation of the die rolls, e.g. not rolling the die properly. We used a wooden table and a 5-row game sheet to keep track of the number of completed die rolls, which allowed us to obtain data on such deliberate manipulations (Figure B3). Limiting the number of rows to 5 per page allowed us to assess the progress during the experiment, whenever participants flipped a page. The wooden table ensured that each die roll was audible. One of the authors noted down the outcomes of each die roll transmitted by the Bluetooth die.

After all milkmen within a neighborhood had completed the experiment, they were invited to fill a post-game questionnaire. Upon completion of this survey, milkmen were paid the sum of earnings plus a show-up fee of INR 200. On average, each milkman earned INR 495 ($\sim$ USD 8).

We also took great care to address the problem of contagion and contamination among milkmen. To mitigate this risk, we conducted the experiment with all milkmen from one dairy neighborhood on a single day and invited all of them at the same time. One of the authors and an assistant monitored their conversations and made sure that they did not discuss the experiment.

Figure B2: Bluetooth die



ID____________ **Date**______________

Roll Number	Die outcome
1	
2	
3	
4	
5	

Figure B3: Game sheet - Die game

References

NSMA, 2011. Executive Summary on National Survey on Milk Adulteration. New Delhi: FSSAI.

HOUSE OF FINANCE

WORKING PAPER SERIES

Recent Issues

No. 133	Markus Behn, Rainer Haselmann, Thomas Kick, Vikrant Vig	**The Political Economy of Bank Bailouts**
No. 132	Rainer Haselmann, David Schoenherr, Vikrant Vig	**Rent-Seeking in Elite Networks**
No. 131	Nicole Branger, Patrick Grüning, Christian Schlag	**Commodities, Financialization, and Heterogeneous Agents**
No. 130	Giuliano Curatola	**Optimal Consumption and Portfolio Choice with Loss Aversion**
No. 129	Giuliano Curatola, Michael Donadelli, Patrick Grüning, Christoph Meinerding	**Investment-Specific Shocks, Business Cycles, and Asset Prices**
No. 128	Giuliano Curatola	**Preference Evolution and the Dynamics of Capital Markets**
No. 127	Helmut Elsinger, Philipp Schmidt-Dengler, Christine Zulehner	**Competition in Treasury Auctions**
No. 126	Carsten Bienz, Karin S. Thorburn, Uwe Walz	**Coinvestment and risk taking in private equity funds**
No. 125	Tobias H. Tröger, Uwe Walz	**Does Say on Pay Matter? Evidence from the German Natural Experiment**
No. 124	Adrian Buss, Bernard Dumas, Raman Uppal, Grigory Vilkov	**The Intended and Unintended Consequences of Financial-Market Regulations: A General Equilibrium Analysis**
No. 123	Marie Lalanne, Paul Seabright	**The Old Boy Network: The Impact of Professional Networks on Remuneration in Top Executive Jobs**
No. 122	Douglas Cumming, Uwe Walz, Jochen Werth	**The Dynamics of Entrepreneurial Careers in High-Tech Ventures: Experience, Education, and Exit**
No. 121	Elia Berdin, Matteo Sottocornola	**Insurance Activities and Systemic Risk**
No. 120	Matthias Heinz, Heiner Schumacher	**Signaling Cooperation**