CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

Discussion Paper No. 2022-03

Devesh Rustagi
Markus Kroell

February 2022

Measuring Honesty and Explaining Adulteration in Naturally Occurring Markets

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Suzanne Robey
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 95 14763
suzanne.robey@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Measuring Honesty and Explaining Adulteration in Naturally Occurring Markets*

Devesh Rustagi† and Markus Kroell‡

Abstract

There is astounding variation in product quality sold in markets even when quality is difficult to ascertain and rules are poorly enforced. We investigate whether sellers differ in innate honesty (incur private cost to provide good quality) and whether this explains the variation in quality. Our study takes place in milk markets in India, where milkmen collude on price, customer rarely switch, and it is difficult to establish reputation. We invite milkmen to take part in a novel behavioral experiment to measure dishonesty. We then measure quality objectively as the percentage of water added to a liter of milk sold to customers. Our results show that dishonest milkmen add significantly more water to milk. Evidence from milk-testing tournament confirms that milk quality is difficult to verify. These results suggest that some sellers are willing to forego monetary gains to provide good quality in return for utility from being honest, even in an environment that encourages cheating.

JEL: C93, D91, O13, Q01

Keywords: Honesty, adulteration, milk markets, asymmetric information, measurement error, India

*We acknowledge financial support from Sustainable Architecture for Finance in Europe (SAFE) under grant number 2014-1228 P2015. We thank Milkotronic Ltd. for providing us with the milk analyzer, and Ms. Maria Arbova for advising us on the machine usage. We especially thank our research assistants, Iqbal Malik, and the NGO Vatavaran for supporting our fieldwork in Delhi. We are also grateful to seminar participants at Goethe University Frankfurt, LMU Micro Workshop, NGO Conference Bad Homburg, Nova School of Business and Economics Lisbon, DIW Berlin, 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. Our special thanks to Andrew Foster, Guido Friebel, Rema Hanna, Rocco Macchiavello, Karthik Muralidaran, and Uwe Walz for providing helpful comments on the paper.

†Corresponding author: Devesh Rustagi, devesh.rustagi@nottingham.ac.uk. University of Nottingham, Department of Economics, Sir Clive Granger Building, University Park, Nottingham - NG7 2RD, United Kingdom. Brown University, Department of Economics, 64 Waterman Street, Providence, RI - 02912, USA. Center for Development Research (ZEF), Bonn University, Genscherallee 3, D-53113 Bonn, Germany.

‡Private sector

I. Introduction

In markets with asymmetric information sellers have a strong incentive to cheat their customers by selling poor quality products at the price of good quality ones. Yet, many markets with such characteristics do not converge to an equilibrium in which all sellers cheat by the maximum possible extent. Rather, we observe large heterogeneity in quality sold even within the same sector, with many sellers refraining from cheating. Understanding why sellers vary in their cheating behavior is important for designing optimal contracts and government interventions. In this paper, we investigate whether sellers vary in innate honesty and whether this variation explains the large heterogeneity in product quality sold in markets.

A prominent view among economists is that individuals cheat rationally, i.e. when the benefits exceed the costs of cheating (Becker 1968). Accordingly, many studies have focused on designing mechanisms, such as monitoring, that increase the costs of cheating (Nagin et al. 2002, Olken 2007, Duflo et al. 2013). Alternatively, individuals may rely on building relationships based on reputation to deter short-term opportunism against future benefits (Baker et al. 1994, McMillan and Woodruff 1999, Banerjee and Duflo 2000, Macchiavello and Morjaria 2015). Recent evidence suggests that individuals may cheat less even when both enforcement and reputation formation are difficult because of psychological costs arising from innate honesty.[1] This may lead to privately costly behaviors, whereby individuals are willing to forego monetary gains from cheating in return for additional utility from honesty.

Such innate honesty is believed to matter in asymmetries affecting health and education because in these sectors providers are socially obliged to serve people (Arrow 1963). However, when it comes to markets, conventional wisdom suggests that sellers are profit-maximizers operating with venal motives. So far, there is no empirical evidence on the role of innate honesty in explaining the variation in product quality sold in markets. We fill this gap by conducting our study in naturally occurring markets for buffalo milk in India.[2]

The milk markets provide an ideal setting for the following reasons. First, milk is a credence good, so it's quality cannot be determined by experience. This gives sellers the incentive to adulterate milk up to a point below which customers cannot detect adulteration, and then sell the adulterated milk at the price of pure milk. Though the law prescribes punishment for adulteration, monitoring is rare and prosecutions never happen. Second, price does not signal milk quality because most milkmen collude on price, which is sustained by dense caste networks. Moreover, the production technology

[1] See Arrow 1963, Fehr et al. 1997, Platteau 2000, Gneezy 2005, Fisman and Miguel 2007, Karthik 2009, Fafchamps 2010, Fischbacher and Foellmi-Heusi 2013, Gachter and Schulz 2015.

[2] These markets account for slightly over 50 percent of the milk sold in India, which itself is the largest producer of milk in the world (see FAO 2013, page 25).

of milkmen does not allow them to distinguish between customers, and everyone is sold the same milk quality, which limits the scope of reputation. Third, customers do not switch across sellers in part because they hold optimistic beliefs over milk quality even though they cannot verify it. Sellers offer credit to long-term customers which further discourages switching. Under these conditions, markets are unlikely to unravel and all sellers are expected to supply the same quality of milk. Moreover, if sellers maximize only profits then this quality is expected to be low and bunched near the threshold beyond which adulteration becomes perceptible. Yet, we find a large variation in milk quality sold in markets. We show that this variation is mainly due to differences in sellers' innate dishonesty.

Milk markets share a range of characteristics with other markets, making the findings from this study of potentially wider interest. They have common attributes with markets for products whose quality cannot be verified by customers and where fraud is common. These range from food products and mechanics to doctors and computer specialists (see Dulleck and Kirschbammer 2006, Dranove 2012, Sheldon 2017).[3] The structure of milk markets shares similarities with other markets in developing countries, especially in sectors with high fixed costs, low capital investment, and strong networks (see Fafchamps 2004, Munshi 2007). A study commissioned by International Livestock Research Institute suggests that informality and non-competitiveness might be a common feature of milk markets in developing countries (see Staal et al. 2008). Also, adulteration of milk is not specific to India alone but is a problem throughout Asia, Sub-Saharan Africa, and Latin America (Liu et al. 2010, Souza et al. 2011, Faraz et al. 2013).

We measure objectively the milk quality sold by sellers (hereafter, milkmen) in actual market transactions. Our assistants pose as occasional customers and purchase a liter of buffalo milk from milkmen in a one-shot interaction and record the price at which the milk was sold. This procedure offers three advantages. First, the one-shot interaction limits the scope of reputation. Second, it allows us to obtain measures of milk quality that are not prone to scrutiny. Milkmen are not aware that they are being studied and hence they cannot alter the milk quality sold to our assistants (see Banerjee et al. 2012). Third, since milkmen do not distinguish between customers, our assistants are sold the same quality that is sold to customers, allowing us to capture adulteration in actual market transactions.

Adulteration of milk happens largely in the form of added water, and the threshold beyond which added water becomes perceptible is reported by milkmen as 30-35 percent. We use an ultrasonic milk machine to determine the precise levels of added water in a liter of milk (error rate $\pm$ 3 percent). We find that the milkmen add water in the range

[3] Examples from food markets, which are closely related to milk markets, include issues of food safety like pesticide residues, ethical production concerns like animal welfare, modes of food production like genetic modification, location via geographic indicators, and better trading conditions via fair trade. None of these attributes can be easily verified by customers.

of 4-37 percent, the average being 18 percent. Since the natural variation of water in milk occurs in very tight intervals (Buettel et al. 2008), the machine captures purposeful adulteration. A second sample of milk obtained for nearly one-half of the milkmen in our study shows strong persistence in adulteration. Since milkmen rear buffaloes themselves and sell milk directly to their customers, we can tie this adulteration to milkmen.

Measuring dishonesty using observational data is likely to be of little value because of confounding with other motives. Behavioral experiments allow researchers to separate dishonesty from confounding motives by controlling for the decision environment. We conduct an experiment in which milkmen have to roll a six-sided die 40 times in private and are paid for each self-reported point. Typically, researchers use these self-reported points to construct measures of dishonesty (see Hanna and Wang 2017). However, such measures are susceptible to measurement errors from confounding dishonesty with the randomness of the device, and also from randomness in behavior because we observe realizations but not the actual tendency for dishonesty. These errors could result in inconsistent estimates and attenuation bias.[4] To mitigate this concern, we conduct our experiment with a Bluetooth-enabled die, which allows us to obtain the actual outcomes of the die rolls. [5] We use this to construct measures of dishonesty that are free from the measurement error that is due to the randomness of the device. Though these measures are still prone to measurement error from the randomness in behavior, this may be small enough for researchers to ignore it (see Harmon 2021). Specifically, we measure dishonesty as the number of over-reported rolls, that is, the difference between self-reported and actual outcomes. We find that 50 percent of the milkmen behaved dishonestly, over-reporting on average in 3.6 rolls (s.d. 6.6) by 7.5 points (s.d. 15.85).

We proceed by examining the association between our measure of dishonesty and adulteration of milk sold to customers. Our purpose here is not to establish causality, but to provide a measure of dishonesty and validate its importance for product quality sold in actual markets. We find a strong positive association between dishonesty and adulteration of milk, which is significant at the 1-percent level. A one standard deviation increase in the number of over-reported rolls is associated with an increase in added water by over 3 percentage points, which is nearly one-half of the standard deviation of the mean added water. Our dishonesty measure explains up to 19 percent of the variation in adulteration,

[4]If the measurement error is classical, it is possible to solve this problem using an instrumental variables (IV) approach. If the measurement error is non-classical, OLS and IV may be used to bound an estimate, yielding respectively lower and upper bounds (Angrist and Kruger 1999).

[5]We designed our instructions carefully. On the one hand, milkmen were not aware that the die is Bluetooth enabled, but on the other hand, we did not lie to them that their behavior in the experiment will never be observed. Our study is not unique in using such a design, as prominent laboratory (Andreoni 1988, Gaechter and Thoeni 2005) and field (Bertrand and Mullainathan 2004, Dizon-Ross et al. 2015, Das et al. 2016) studies also withheld information to study the effect of surprise re-starts, sorting by types, discrimination, and cheating respectively. As in previous studies, such a design allows us to obtain otherwise inaccessible data – actual outcomes of die rolls – without harming the participants (see Glennerster and Powers 2014 and Cason and Wu 2019).

the largest of all covariates. The strong positive association also holds when we use data from the second sample. These magnitudes are economically substantial as the additional monetary gain from dishonesty amount to eight percent of a milkman's monthly income. These results indicate that honest milkmen are willing to forego substantial monetary gains in return for additional utility they get from being honest.

We show that the above results are robust in many ways. First, we control for important milkman specific characteristics, as well as livestock related input factors that are likely to affect milk quality. In doing so, we pay particular attention to production costs (buffalo herd size, lactation, and feeding costs), economic variables (wealth, family size, outside option), and other relevant behaviors (altruism, spite, reciprocity), education, caste, and religiosity. Second, we account for unobserved differences across different markets from where milkmen operate their dairies via market fixed effects. Third, we also account for fixed effects for the assistants who bought the milk. Finally, we control for price even though there is little variation in it because of collusion. These covariates are meaningful as their inclusion leads to a large jump in R-squared. Following this insight, we find that selection on unobservables would have to be six times larger than selection on observables to explain away our results, which seems unlikely.

Our motivation to use a Bluetooth-enabled die to measure dishonesty was to mitigate the attenuation bias from measurement errors. We find that our approach is indeed warranted because the error in self-reported measures turns out to be large. While it is possible to use a split-sample instrument to address this concern, finding valid instruments may be difficult. In our context, such an instrument does turn out to be invalid as the measurement error in the first set of rolls is correlated with the second set of rolls. These results highlight the relevance of a Bluetooth-enabled die in measuring dishonesty.

We now turn to presenting evidence from an incentivized milk-testing tournament to confirm that milk is a credence good. In the tournament, we asked milkmen and customers separately to predict the quality of five different milk samples containing 0, 100, 200, 300, and 400 ml of added water in a liter of milk. The first four samples contained added water below the threshold at which adulteration is perceptible, whereas the fifth one was close to this threshold. While analyzing the predicted level of added water, we find that customers and milkmen cannot distinguish between the different samples and predict each sample as having the same level of added water. However, while analyzing the rank of each sample, we find that customers and milkmen rate the most diluted sample slightly worse than the pure sample. These results show that milk quality is difficult to verify when adulteration levels are below the threshold. This also explains why we hardly observe added water above the threshold in our sample.

This paper contributes towards understanding the large variation in product quality observed in markets in developing countries. Previous studies focus on the role of mechanisms that affect pecuniary costs of cheating. For instance, Macchiavello and Morjaria

(2015) provide evidence on the role of reputation in the context of Kenyan rose export sector. Our study adds to this literature by highlighting the importance of innate honesty in markets where monitoring and reputation are less likely to play a role. These results can be interpreted as suggesting that honest sellers are willing to forego monetary gains from cheating to maintain their sense of morality. While previous studies show the importance of dishonesty for other outcomes (see Dai et al. 2017, Hanna and Wang 2017, Cohn and Marechal 2018, Barfort et al. 2019), few papers have shown this for behaviors that are as directly economically relevant as the one studied here - producers choice of product quality.

Our paper also contributes to the literature on measurement of honesty at the individual level. Previous studies rely on self-reported outcomes that are susceptible to errors in variables. Our paper mitigates this problem by using a Bluetooth-based technology to obtain actual outcome of the die rolls that we use to construct measures of dishonesty that are less prone to errors in variables.

This paper also connects to the literature that attempts to measure corruption and cheating outcomes objectively rather than relying on perceptions. For example, Olken (2007) uses engineers' estimation of costs based on materials used in road construction to estimate corruption. Hanna and Wang (2017) randomly visit several dispensaries multiple times to assess absence from duty. We use a machine to obtain precise levels of added water in milk sold in actual market transactions.

The rest of the paper is organized as follows. Section II describes milk markets in India. Section III describes data on milk quality sold in milk markets and the behavioral experiment to measure dishonesty. Section IV presents our empirical strategy and main results. Section V presents results from the milk-testing tournament. Section VI offers concluding remarks.

II. Field Setting

This section provides background information on milk markets in Delhi, where we conducted our study. It relies on secondary data, as well as primary data we collected via in-depth surveys and group discussions with milkmen and customers.

There are 10 markets for buffalo milk on the outskirts of Delhi, which are, on average, 12 km apart from each other. Each market is a neighborhood from where several milkmen operate their dairy. The neighborhoods were designated exclusively for dairy farming in the 1970s under the local zoning law to move milkmen out from the central to the peripheral areas of the city (DDA 1985). Each milkman was allotted a plot in one of the neighborhoods to construct his dairy. The neighborhoods were organized into blocks of 6-8 adjacent dairies of the same size (see Figure 1). Usually, a dairy is a large room with two rows of buffaloes, separated by a drain that disposes water and waste out of the dairy. A typical dairy is poorly ventilated, tightly packed, and smells strongly of livestock waste. A large majority of the milkmen own a single dairy; we did not come across any milkman with dairies in different neighborhoods.

Our focus is on neighborhoods where milkmen reside locally and operate the dairies themselves without any outside labor. This allows us to match at the milkman level adulteration in milk with dishonesty in the experiment. The ground floor of the dairy houses buffaloes, whereas the milkmen and his family members reside on upper floors.

Figure 1: Dairy Structure and Operation

Notes. The left panel shows high density of dairies within a neighborhood. The right panel shows customers lining in a short queue outside a dairy to buy milk.

Dairy operation.– The milkmen sell fresh non-pasteurized buffalo milk from the doorsteps of their dairy daily in the morning (6-8 am) and in the evening (4-6 pm) directly to customers (without any intermediaries). The milk is sold loose, that is, it is non-packaged and non-bottled.

There is no mechanization and buffaloes are milked manually inside the dairy before the customers arrive. Milking buffaloes in front of the customers is not an option because it is logistically demanding. Buffaloes are washed, fed, and prepared before milking. A calf is brought to induce milk flow and hind legs are tied to prevent kicking. Once begun, milking cannot be stopped to allow for each customer to observe. especially given that milkmen have on average 19 buffaloes. Also, the dairies don't have the space to accommodate the customers. The dairies are already packed with buffaloes and smell strongly of livestock waste. For these reasons, only milkmen and family members enter the dairy.

After milking the buffaloes, the milk is pooled in a large container or a bucket kept outside the dairy from which it is then sold to customers. The customers line up in a short queue (see Figure 1) and typically buy 1-2 liters of milk per day.[6] All customers are sold

[6]During our field visits, we never saw customers waiting in a long line.

the same milk quality and are also charged the same price. Milkmen do not invest into branding or advertising. None of the dairies have a name or any kind of marker on their facade. Investments in management practices that could allow milkmen to differentiate along hygiene, waste management, and animal welfare are completely lacking.

Milk production is driven in part by storage possibilities, capacity constraints, and lactation cycle of buffaloes. Milk is highly perishable and it must be sold within a few hours of milking buffaloes. In the absence of cold storage and product diversification, one way to expand is to steal the customers of other milkmen. However, this is highly implausible, as it would result in conflict between milkmen which could get amplified because of strong social-ties (from instance from caste). Since each plot can accommodate only a certain number of buffaloes, milkmen have the option to expand by taking over the plot of another milkman. This is also unlikely because milkmen who rent their plot could either face eviction for not operating a dairy or would have to bear costs for land use change. We did not come across any milkman who was renting his plot to other milkmen. Milkmen could also expand their production by adding more water on certain days. However, a strong positive association in adulteration across the two rounds of data that we have for some milkmen suggests that adulteration is stable (See Figure 2).

The milkmen have two types of customers: repeat and occasional. The majority of customers are in a repeated interaction, with 80 percent buying milk from the same milkman for at least 2 years. The repeat customers buy milk daily on credit and pay at the end of the month. There are no written contracts; instead, each milkman maintains a book with name, address and phone number of repeat customers.[7] The occasional customers buy milk on an irregular basis and pay in cash at the time of purchase. Because milkmen sell milk from the same container they do not differentiate between different types of customers – all are sold the same milk quality at the same price.

Customers.–The customers live in residential areas that are within the walking distance of the dairy neighborhoods. New customers are rare, at the most 1-2 per neighborhood per month. The customers live in residential areas that are within the walking distance of the dairy neighborhoods. A new customers' choice of milkman follows the recommendation of existing customers who live in close proximity.

The customers expect milkmen to supply pure quality milk because it is nutritious. Over 90 percent of the milkmen reported selling pure buffalo milk and stated further that "if customers wanted diluted milk, they could add water themselves". This means milkmen are aware that customers expect high quality milk. Though customers cannot verify quality, they still believe that their own milkman sells good quality, giving it a score

[7] The milkmen were reluctant to give us the names and address of their customers because this information is confidential. In addition, we were told that if repeat customers fail to show up on a given day, they are expected to inform in advance and are not charged for that specific day.

of 8 out of 10 points (std. 1.20).[8] Milkmen also stated that customers complain rarely about the milk quality sold by them. Other factors like timeliness of milkmen and getting priority on festival days appear to be of less importance. This is in part because it is the customers who go to milkmen and not the other way round. In addition, during festival days, customers meet high demand by buying packaged milk.

Given that the customers are satisfied with milk quality, it is not surprising that they rarely switch between milkmen. In the survey, customers reported buying milk from the same milkman, on average, for over 14 years. This is also because of credit that milkmen offer to long-term customers. Milkmen also discourage switching to avoid conflict with each other. Moreover, to switch, customers need to seek out information about other milkmen from other customers. However, to get enough information they must speak to several of that milkman's customers. As the number of milkmen increases it becomes increasingly difficult to collect this information, making consumers less likely to switch (see Satterthwaite 1979, Dranove 2012).

Market structure. – Milk markets operate from different neighborhoods and each resembles a local oligopoly. Although there is differentiation in milk quality from added water, quality is not easily verifiable. There are high barriers to market entry and exit. The zoning law restricts milkmen to operate from designated neighborhoods only. New milkmen can enter the market only if they can acquire a plot in one of these neighborhoods. Since no new plots are available, market entry is almost impossible. Also, since dairy operation in these areas is still largely a caste-based occupation, there is little interest in large-scale capital investment. Similarly, the costs of exiting the market are also high because it requires milkmen to change the land use from agricultural to residential /commercial. In some neighborhoods such a change is not possible and milkmen would be evicted from their house if they decided to quit their dairy-based occupation. In other neighborhoods, this entails a fixed fee of INR 100,000 – 400,000, plus a fee of INR 15,000 – 25,000 per m^2 for the plot, and additional costs for floor to area ratio by INR 3000 – 8000 per m^2. The charges are twice as high if the conversion of land is for commercial use (DDA 2006, DDA 2018). Since these costs can easily add up to millions of rupees, exit is rarely heard of.

70 percent of the milkmen collude on price within a neighborhood. The variation in price is therefore very small (s.d. 4.278). This collusion is mostly, if not exclusively, along caste lines. For instance, 88 percent of the milkmen who belong to the *Gujjar* and *Jat* caste groups charge the same price. The price is set in a group meeting and it is also raised collectively by a fixed amount every year. The milkmen who collude on price are

[8]For instance, a customer remarked that she has been in a repeated interaction with a milkman for 25 years. She vouched for the pure quality of the milk that her milkman supplies and this as the reason behind the long-term relationship. However, when we tested the milk on two different occasions, both times the milk turned out to be containing nearly 30 percent water.

also the ones who dominate the milk market, producing together over twice the volume of milk sold than milkmen who are not part of this collusion. Milkmen collude on price because it is easy to verify than quality (see Section V).

Several characteristics of milk markets ensure that price collusion is self-enforcing. First, milkmen live and operate from neighborhoods that are stable in composition because of high barriers to market entry and exit. This means that the interaction among milkmen can be considered as infinitely repeated. Second, high density of dairies (Figure 1) and the possibility of interaction with customers of different milkmen make it easier to monitor price deviations. Third, most milkmen belong to the same tight-knit caste groups, which makes deviations costly with a possibility of social boycott. Such strong punishments might outweigh potential short-term gains from deviation. Fourth, milkmen strongly discourage customers from switching precisely to dissuade deviations. Fifth, daily fluctuations in demand and supply of milk are rare and occur only during the main festival days of Hindus (Diwali) and Muslims (Eid), when milk is in high demand. Sixth, the scope of deviation from collusion via other avenues is unlikely because milkmen sell only liquid milk and no other ancillary milk products. Seventh, since customers buy usually 1-2 liters of milk, buyer power is unlikely to break this collusion.

Price collusion means that honest milkmen are unable to signal quality via price. Others means of signaling like buying own machine and letting customers test milk directly on site are unlikely because the machine is expensive and beyond the means of individual milkmen.[9] As mentioned above, milking buffaloes in front of the customers is also not an option. This means honest milkmen have lower revenues than dishonest milkmen. This in itself does not drive honest milkmen out of the market or cause milk markets to unravel. This is because of several reasons: collusion allows milkmen to charge a higher price, milkmen adulterate below the threshold at which quality becomes verifiable, customers rarely switch and believe that milkmen supply good quality, and high exit costs.

Milkmen face little competition from the milk sold by cooperatives or private sector, which predominantly sell packaged and pasteurized cow milk. The buffalo milk differs from cow milk in having higher content of fat, protein, and solids-not-fat (SNF) and is thus much more nutritious (Menard et al 2010).[10]

Milkmen characteristics. – All milkmen are male and inherited milk business from their father. 67 percent have high school education and have on average seven family members. Most milkmen are from the tightly knit *Gujjar* and *Jat* caste groups. Milk sale is the main source of income for 58 percent of the milkmen who don't have any household member working in a non-dairy sector. Most milkmen are from states neighboring Delhi; the

[9] The machine costs around INR 88,000. We believe it is highly unlikely that honest milkmen are individually going to invest in a machine worth the price of over 1500 liters of milk.

[10] The predominance of cow milk in the formal sector could be because many cooperatives that supply milk to the formal sector keep cows. On average, the price of a cow is half that of a buffalo.

largest share is from the state of Uttar Pradesh.

III. Data

We collect data using five different sources: (a) milk purchase via mystery shopping, that is, via assistants not known to the milkmen to assess milk quality sold in markets; (b) a behavioral experiment to measure innate honesty at the individual milkman level, (c) milk-testing tournaments with milkmen and customers to assess verifiability of milk quality; (d) several rounds of community and household surveys with milkmen and customers. These surveys cover a variety of questions including socio-demographic characteristics, dairy operation, contractual practices, interaction with customers, pricing of milk, livestock related input factors, choice of milkmen, perception of milk quality, switching, and duration of repeated interaction; and (e) a behavioral experiment to measure other relevant behaviors, such as altruism, spite, equality and selfishness. The behavioral experiments were conducted in November 2014. The milk samples were purchased a month later in December 2014. The tournaments were conducted with milkmen in March 2015 and with customers in August 2018. The community and household surveys were implemented during this period. The final round of consumer survey was conducted in 2021. We describe below our sample, data on milk quality, and behavioral measures of honesty. Data from the milk-testing tournament are described in Section V.

III.A. Sample construction

Our sample comprises milkmen from six dairy neighborhoods, of which five are from Delhi and the sixth one is from an adjoining city in Uttar Pradesh.[11] These neighborhoods cover all regions of Delhi: East, North East, North West, South, South West and South East. Within these neighborhoods, we focus on small milkmen. The detailed maps together with the baseline survey that we conducted suggest that there are roughly 160 small milkmen residing in these neighborhoods. The small milkmen conduct the dairy operations themselves including selling milk, feeding, washing, and milking buffaloes. We have 72 milkmen from six neighborhoods in our main sample.[12] In addition, 98

[11] Of the ten dairy neighborhoods in Delhi, we excluded two neighborhoods because of their small size. Another three were excluded because they comprise exclusively of large milkmen who do not operate the dairy themselves but employ outside labor to do that. This makes it difficult to identify the individual who adulterates milk and hence mapping the association between dishonesty and adulteration.

[12] Actually 84 milkmen took part in the behavioral experiment to measure dishonesty. We could not collect milk sample from nine milkmen and for another three milkmen the information from the Bluetooth die is missing. Excluding these 12 observations leaves us with a final sample of 72 milkmen. We verify that these 12 milkmen do not differ from the rest in their self-reported outcomes in the experiment or their socio-demographic characteristics (see Table A1). It was not possible to include all milkmen in our study. The use of the Bluetooth die meant that every milkman required our personal attention. Therefore, we could not run the experiment with multiple milkmen at the same time, as would have been possible with a normal die. Also, milkmen have limited free time in which we had to accommodate our study. This

milkmen from four neighborhoods and 41 customers living in the proximity of two of these neighborhoods took part in the milk-testing tournament.

There is a possibility that our main sample is prone to selection as participation in the experiment was voluntary. To mitigate this concern, we compare milkmen in our sample to milkmen who are not in our sample across two key dimensions – quality (added water in percent in a liter of milk) and price for a liter of milk. Table A.2 reports the results and shows that the milkmen in and out of our sample are very similar with respect to both added water and price. A regression-based test in columns 3 confirms this and shows that the differences are small and statistically insignificant. This is because milkmen who did not take part in our study are also small milkmen who operate the dairy themselves and reside in the same neighborhoods as milkmen who took part in the study.

III.B. Milk quality in milk markets

Generally speaking, obtaining reliable measures of product quality is difficult. Much of the previous work is based on perceptions that can be attributed to a multitude of factors. As mentioned before, our approach allows us to obtain objective and reliable measures of milk quality sold in markets. However, because milkmen may add besides water a variety of other adulterants, measures relying exclusively on added water may underestimate the extent of adulteration. Accordingly, we first verify in a pilot study that added water is the only adulterant. We purchased a liter of milk from 15-20 milkmen from each of the six neighborhoods in our study and then split each milk sample into two parts. One part was tested for a variety of adulterants including water by a professional food-testing laboratory and the other part was tested for added water using an ultrasonic milk analyzer (see Appendix A). Besides confirming that added water is the only adulterant, the laboratory results are also strongly correlated with the machine result ($r = 0.93$). Accordingly, we use added water in a liter of milk as our measure of milk quality (adulteration) in markets.

Matching adulteration with measures of dishonesty at the individual level is demanding because it requires having the precise location of milkmen. Since the dairies typically do not have plot numbers or other identifiers such as a name, this is a daunting task. Therefore, we prepared detailed maps listing each milkman within a neighborhood and marked carefully the ones who took part in the experiment as well as the ones who did not (Figure A.1).

We then hired assistants unknown to the milkmen to pose as occasional customers and buy a liter of milk in a one-shot interaction. Because milkmen did not know that the same team is behind the behavioral experiment and milk purchase, they could not alter the milk

together with our strategy of conducting the behavioral experiment within a neighborhood on a single day to prevent contagion restricted the sample size further. In the experimental pilot studies that we conducted, a session with 15 participants lasted up to three hours including the post-experimental survey and payments.

quality sold to our assistants; they were sold the same quality as other customers. Also, the milk samples were collected a month after the behavioral experiment was conducted and therefore we doubt that the experiment itself affected the milk quality sold to our assistants. The assistants noted the price at which the milk was purchased. All samples from a neighborhood were obtained on the same day.

The milk samples were tested for adulteration via an ultrasonic milk analyzer. The analyzer uses the freezing point of buffalo milk to calculate the percentage of added water in a liter of milk and has an error rate of $\pm$ 3 percent. The freezing point of milk lies below the freezing point of water. As more water is added to milk, its freezing point moves closer to that of water (Advanced Instruments 1995).[13] We find that on average 18 percent or 180 ml of every liter of milk sold by milkmen is actually water. Figure 2 shows large variation in the added water in milk, which ranges from as low as 4 percent to as high as 37 percent.

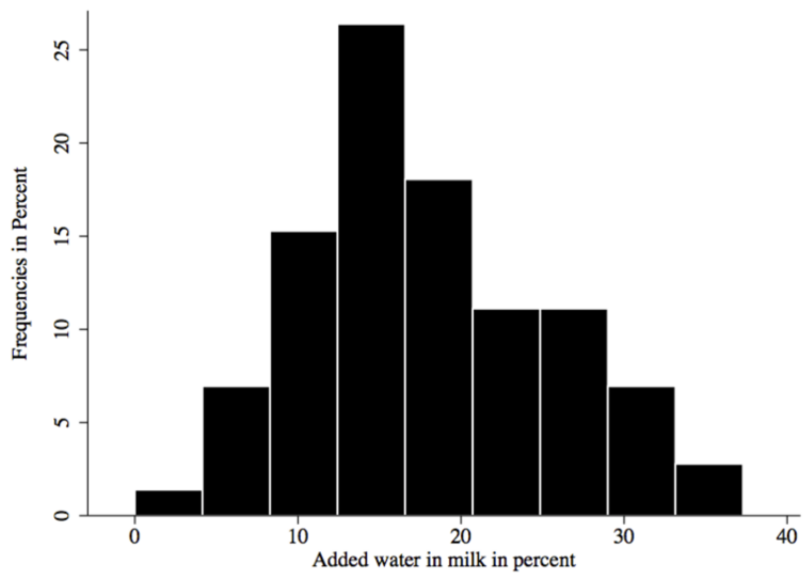


Figure 2: Histogram of added water in a liter of milk (ml)

We consider the above variation in milk quality as reflecting adulteration by milkmen for the following reasons. First, this variation cannot be due to natural variation of water in milk. The freezing point of milk is a biological constant that varies within very tight intervals (Buettel et al. 2008). Therefore, the wide variation in added water that we observe cannot be due to the natural variation of water in milk. Second, the variation in milk quality cannot be due to livestock variety because over 90 percent of the buffaloes are

[13] 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*.

of the same breed. Third, the variation cannot be due to weather fluctuations because we collected milk samples from each neighborhood in the third week of December. Fourth, the variation in milk quality is also not due to fluctuations in demand and supply, as these are rare outside the festival periods and we did not conduct our study during festivals.

If milkmen add water only on some days then the above measure of milk quality may not be reliable.[14] To this end, we also collected a second round of milk sample for 29 milkmen in our study after 10 days of collecting the first sample. These milkmen were selected randomly. Table A.3 shows that they are comparable in characteristics including adulteration and dishonesty to other milkmen. Figure 3 shows that there is a strong positive and significant association in milk quality across the two samples. In a regression of the second sample on the first sample after controlling for covariates and fixed effects, the coefficient on the first sample turns out to be 0.91. This means if the added water in the first sample is 10 percent, the predicted value of added water in the second sample is 9 percent, thereby suggesting stability in adulteration.

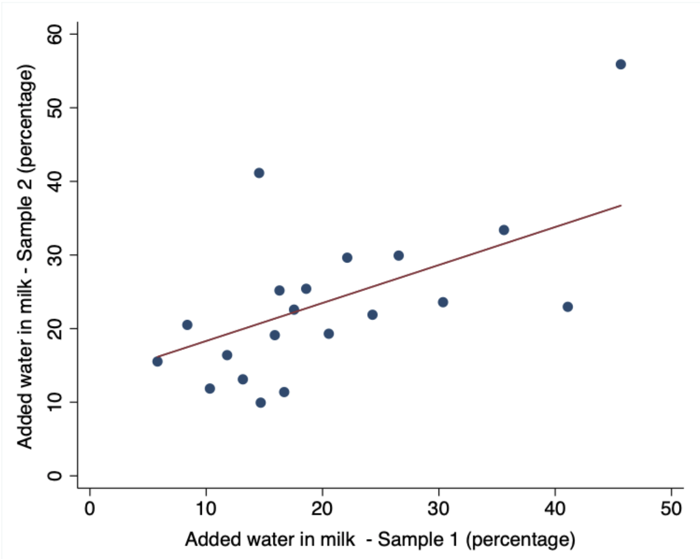


Figure 3: Scatter plot of added water in milk in the second (y-axis) and the first sample (x-axis)

From Figure 2 we see that very few milkmen did not add any water to the milk. It could be that when the milkmen add water, they are not comparing with zero percentage, but instead with some relative percentage, which may vary across milkmen. We address this concern as follows. First, we asked milkmen in surveys the milk quality their customers demand and the milk quality they sell. Over 90 percent of the milkmen stated that their customers exclusively demand pure milk and an equally high percentage reported selling

[14]See for instance Athey and Bagwell (2001) who show that firms use market share exchanges to achieve optimal collusion. However, this mechanism is not at work in milk markets because a large proportion of customers are in a repeated interaction and do not switch from one milkman to another.

only pure buffalo milk. These findings clearly suggest that pure milk is the standard that customers expect from milkmen. Second, from the location of each milkman, we construct an alternative dependent variable that captures added water relative to the nearest neighbor. We subtract from each milkman's added water the added water by the nearest milkman in the neighborhood in our sample. This of course assumes that milk quality is observable and that milkmen are aware of the quality sold by their neighbors. Although we doubt that these assumptions hold (given the results of the milk-testing tournament), we use this alternative dependent variable while conducting robustness checks.

III.C. Behavioral experiment to measure honesty

Measuring honesty is challenging because of potential confounds arising from opportunities for repeated interaction and reputation formation. Behavioral experiments allow researchers to exert control over these factors and obtain cleaner measures. Our experimental strategy builds on games of chance, which involve self-reporting of outcomes of random events.[15] Individuals have to roll a die or flip a coin in private and then self-report its outcome. The payoffs depend entirely on 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 innate honesty.

The experimenter does not observe the actual outcomes of random events, so dishonest reporting can only be inferred by comparing self-reported outcomes with the corresponding theoretical probability distribution of the random event. Based on this inference, several measures of dishonesty can be constructed that take into consideration the likelihood and the intensity of misreporting (see Harmon 2021). However, measures based on self-reported outcomes are prone to errors in variables because of two reasons. First, they confound dishonesty with the randomness of the die rolls. This means some individuals might be erroneously classified as more dishonest because their self-reported outcomes surpass the expected outcome by pure chance, others might be classified as more honest despite severe over-reporting, if the self-reported outcome falls below the expected outcome. Second, there is randomness in behavior, as researchers observe only the realizations and not the actual tendency for dishonesty. These concerns imply that in a regression of adulteration on dishonesty, a correlation between self-reported measures and the composite error term would result in inconsistent estimates that are prone to attenuation bias.

These errors can be somewhat mitigated by increasing the level of significance and the number of die rolls, but finding the appropriate number of repetitions per individual is difficult due to both statistical and pragmatic reasons. One the one hand, one has to

[15]Rosenbaum et al. (2014) provide an overview of experimental approaches to measure dishonesty.

correctly forecast the effect size, i.e. the degree of misreporting for a given number of repetitions. On the other hand, one has to consider issues related to the implementation of the experiment (e.g. amount of time, fatigue, and tediousness of the task, which increase with the number of repetitions). Alternatively, it may be possible to use a split-sample instrument to address this concern, but finding valid instruments is difficult. We show in Section IV, that the measures based on self-reporting are indeed prone to large measurement errors and the instrumental variables approach may have a limited role in resolving this problem.

Experimental design.– Given the above concerns with self-reported measures, we conduct our experiment with a Bluetooth-enabled die, which transmits the actual outcome of each roll to the relevant app on a smartphone. We then compare self-reported outcomes with the actual outcomes of the die rolled by each milkman to construct our measures of dishonesty. This allows us to bypass the measurement error from the randomness of the devise and obtain measures of dishonesty that are prone to measurement error only from the randomness in behavior. Later, we present results suggesting that this error is small and can be ignored.

In the experiment, milkmen have to roll a six-sided die 40 times and then self-report the outcome of each roll by striking out the appropriate number of 2 Indian Rupee (INR) coins on a game sheet (Figure A.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.[16] Instructions were neutral and did not encourage dishonesty but explicitly stated to roll the die and report the outcome (see Appendix B for instructions). The responses in the post-game survey confirm that milkmen were aware of the possibility to increase their payoffs by over-reporting the actual outcomes.

While conducting the experiment, we took great care to ensure that every milkman understood the game, carried out the experiment as outlined (e.g. rolled the die) and that contagion across participants was minimized. We gave detailed instructions and examples of the die game that were tested and polished in four pilot studies. In addition, each milkman had to answer three control questions correctly before taking part in the experiment. We outline these procedures in detail in Appendix B.

The experiments were conducted within the neighborhood where the milkmen reside. 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 295 plus a show-up fee of INR 200.

[16]We used higher stakes than Hanna and Wang (2017) who paid INR 1 for every self-reported point to nurses working for a state government in India because milkmen earn much more than nurses. The use of higher stakes does not by itself encourage dishonesty as the self-reported sum that we observe falls between nurses and students from India.

Measures of dishonesty.– 36 out of 72 milkmen in our sample (50 percent) are dishonest as they self-reported higher than the actual outcome. We develop different measures of dishonesty that 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 over all die rolls the number of times 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. On average, milkmen over-reported in 3.6 rolls by nearly 7.5 points. Figure 4 shows that there is positive association between dishonesty measured as the number of over-reported rolls and added water in a liter of milk.

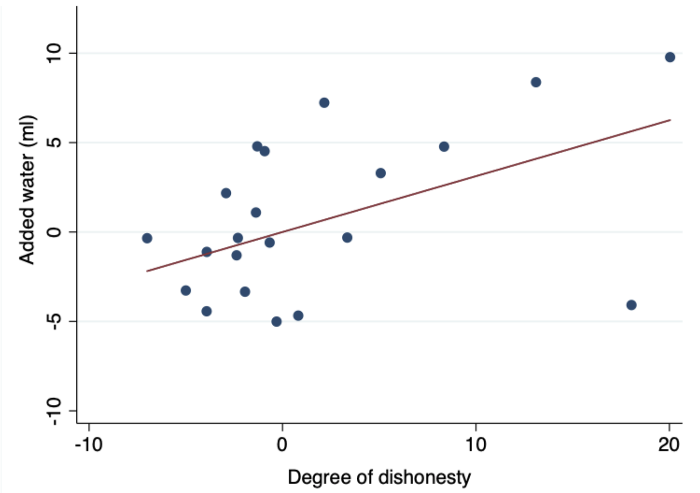


Figure 4: Scatter plot of added water in milk (ml) and degree of dishonesty

Notes. Degree of dishonesty is measured as the number of over-reported rolls in the experiment. The scatter plot accounts for neighborhood fixed effects.

The use of a Bluetooth die itself could have biased our measures of dishonesty. To this end, we conducted an additional experiment in which half of the participants were randomly selected to play the game with a Bluetooth die and the other half with a standard die that can be purchased in local markets. We find that there is no difference in the reporting behavior across these two groups of participants. The mean for the group that used the Bluetooth die is 153.49 points, whereas it is 158.88 points for the group that used the regular die. The difference turns out to be statistically insignificant (p-value= 0.34). A Kolmogorov-Smirnov test further suggests that the two distributions are also not significantly different from each other (p-value= 0.70).

III.D. Behavioral experiment to measure other behaviors

We used dictator games that follow Charness and Rabin (2002) and Chen and Li (2009) to measure equality, selfishness, altruism, and spite. The games were conducted with 48 of the 72 milkmen from five neighborhoods (25 honest and 23 dishonest). Milkmen had to take three decisions introduced in a randomized order. In each decision, they had to choose between two allocations, of which one always yielded equal payoffs (see Table 1). In the first decision, efficiency was held fixed, so if a milkman chose allocation 2 over allocation 1, he displayed altruism. In the second decision, efficiency was also held fixed, so if a milkman chose allocation 2 over allocation 1, he displayed selfish behavior. Finally, in the third decision, if a milkman chose allocation 2 over allocation 1, he displayed spiteful behavior.

Payoff	Allocation 1	Allocation 2
	Decision 1 – Altruism	
Own	200	180
Other	200	220
	Decision 2 – Selfish	
Own	200	250
Other	200	150
	Decision 3 – Spiteful	
Own	200	200
Other	200	180

Table 1: Experiment to elicit behavioral types

We find the following patterns: a) 14 milkmen (29%) always choose allocation 1 in the three decisions (equality); b) 3 milkmen (6%) chose allocation 2 in the first decision, but allocation 1 in the other two decisions (altruists); c) 18 milkmen (38%) chose allocation 2 in the second decision, but allocation 1 in the other two decisions (selfish); d) 8 milkmen (17%) chose allocation 2 in the third decision, but allocation 1 in the other two decisions (spiteful); and e) 5 milkmen (10%) chose conflicting allocations (other).

IV. Dishonesty and Adulteration in Milk Markets

We investigate the association between dishonesty and adulteration in milk markets. Our purpose here is not to establish causality. Rather, we aim to validate the importance of our measure of dishonesty by showing that it explains the extent of adulteration in actual milk markets.[17] We achieve this by using the following OLS specification:

$$y_{in} = \beta_0 + \beta_1 Dishonesty_{in} + \mathbf{X}_{in}\beta_2 + \mathbf{L}_{in}\beta_3 + \alpha_n + \alpha_a + \epsilon_{in} \quad (1)$$

[17] In this sense, our approach mirrors that of Bloom and van Reenen (2007) who measure and validate management practices by showing its importance for firm outcomes.

where y_{in} is the adulteration of milk, measured as the percentage of water added in a liter of milk sold by milkmen i in neighborhood n. Dishonesty is the extent to which a milkman cheats in the die game. In our main specification, it is measured as the number of over-reported rolls. Subsequently, we also present results using other measures of dishonesty, such as the number of over-reported points. **X** is a vector of milkmen specific factors including number of years a dairy is in operation, price charged for a liter of milk, age, education, caste, and religiosity. **L** is a vector of livestock specific factors including number of buffaloes and an indicator for buffaloes in the first phase of lactation. Table 2 presents the definition and summary statistics of these variables. α_n and α_a are fixed effects for the neighborhood where a milkman resides and the assistant who bought the milk, and ϵ_{in} is an in error term. The coefficient of interest is β_1, which captures the association between dishonesty and adulteration. We expect this coefficient to be positive, that is, adulteration increases with dishonesty. We report robust standard errors and while conducting robustness check cluster these on the neighborhood to account for potential correlation of errors within a neighborhood. Since the number of neighborhoods is small, we follow Cameron et al. (2008) and use the wild bootstrap procedure.

A major concern in the estimation of equation (1) is that dishonesty is not randomly assigned. Given the cross-sectional nature of our data, omitted variables bias poses a serious concern. Below we present an overview of several strategies that we use to mitigate this concern and discuss these in detail while presenting the results.

First, we include a rich set of control variables that are hypothesized to matter for milk quality. To the extent differences in milk quality are due to neighborhood specific factors, our specification controls for neighborhoods fixed effects. We also include fixed effects for the assistants who bought milk to alleviate any potential concern arising from them. Second, we carry out a balance test to show that dishonesty is uncorrelated with a variety of covariates. When we regress one by one each covariate on our measure of dishonesty, the coefficients turn out to be small in magnitude and are also statistically insignificant (see columns 3-4, Table 2).[18] We then proceed by regressing dishonesty on main control variables and report the coefficient on each covariate in Table 3. The coefficients on control variables are also small and individually as well as jointly statistically insignificant (p-value = 0.59). These results remain unchanged when we additionally control for neighborhood fixed effects.[19] Third, following Oster (2019), we gauge how large the selection on unobservables would have to be to explain away our results. We also conduct a test based on randomization inference with 5000 replications. Fourth, while conducting robustness checks, we control for additional variables that might be confounded with dishonesty. This includes different measures of economic prosperity (land holding, car ownership, outside

[18]The only exception is the coefficient on family size which is weakly statistically significant at the 10-percent level. The magnitude of the coefficient is very small.

[19]The results do not change when we conduct the same exercise with other covariates from panel D of Table 2.

option, and family size), experimental measures of altruism and other relevant behaviors, and production costs, state of origin of milkmen, behavior of neighbors in the die game, and role of luck in the die game. Lastly, we also use an alternative dependent variable whereby we measure adulteration as added water by a milkmen over and above added water by the nearest neighbor.

IV.A. Main Result

Table 4 presents results on the association of dishonesty with adulteration of milk sold in markets. Column 1 is without any controls and shows that dishonesty has a positive coefficient (0.291), which is statistically significant at the 10-percent level. We introduce in column 2 fixed effects for the neighborhood where a milkman resides and also operates his dairy. The coefficient on dishonesty rises slightly and its standard error declines, such that it is now statistically significant at the 5-percent level. The neighborhood fixed effects are jointly statistically significant at the 1-percent level and their inclusion leads to an increase in the R-squared to 0.17. In column 3, we include milkmen and livestock specific factors together with assistant fixed effects. The coefficient on dishonesty rises to 0.49 and is now statistically significant at the 1-percent level. The control variables are powerful predictors of adulteration in milk markets, as their inclusion leads to a jump in the R-squared from 0.24 in column 2 to 0.50 in column 3; they are also jointly statistically significant (p-value < 0.001).

In column 4, we control for the price at which the milk was purchased. The coefficient on dishonesty remains robust in magnitude and significance, which implies that dishonesty matters over and above price. The coefficient on price turns out to be small in magnitude and is also statistically insignificant. It's inclusion results only in a small increase in R-squared by less than two percent. This result is expected because most milkmen collude on price resulting in little variation within a neighborhood. To have tighter control over price, we consider a sub-sample of only those milkmen who charge the same price. This includes 40 milkmen (56 percent of the sample) who charge INR 60 for a liter of milk. Column 5 reports the results and shows that the coefficient on dishonesty remains nearly unchanged in magnitude and significance. Most of the control variables also retain their magnitude even though the sample now drops by almost half. This clearly shows that our results are not being driven by price.

Among the covariates, the caste dummy and the number of buffaloes enter positively and are statistically significant at the 5 and 1-percent level, respectively. The coefficients on religiosity and lactation period are negative and statistically significant at the 10 and 5-percent level, respectively. The negative coefficient on lactation dummy is expected because during the first half of the lactation period (roughly 3 months) the buffalo milk contains less fat and is less dense (FAO 2013), which could discourage milkmen from adding further water to milk.

We are worried that the above results could be due to some outliers. To counter this concern, we drop in column 6 four most influential observations in the sample (of which two pull the coefficients upwards and the other two downwards).[20] The coefficient on dishonesty retains its magnitude as well as significance. However, the coefficient on price declines in magnitude by almost one-half to -0.163 (s.e. 0.195) and remains statistically insignificant. In fact, the coefficient on price drops even if we exclude just one influential observation (coef. 0.179, s.e. 0.176). A related concern is that our results are driven by a specific neighborhood. We test for this in Table A.4 by documenting changes in the coefficients on dishonesty and price from dropping one neighborhood at a time. The coefficient on dishonesty is over 0.40 in magnitude in all except column 5, where it drops to 0.30, but remains statistically significant at the 5-percent level. In contrast, the coefficient on price fluctuates from column to column and is always statistically insignificant. Together, these results suggest that the association of dishonesty with adulteration is not driven by influential observations and outliers.

Overall, our estimates suggest that milkmen add, on average, 0.475 percentage points more water to milk per over-reported roll (column 4). 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 3.14 percentage points, which is one-sixth of the mean level of added water. When we decompose the R-squared, it turns out that dishonesty explains the largest variation in adulteration in milk markets (19 percent), followed by control variables (45 percent) and fixed effects (36 percent). Following this insight, we find that selection on unobservables would have to be six times larger than selection on observables to explain away our results, which seems unlikely. Our results are further confirmed when we implement a randomization inference test (p-value < 0.001).

Our estimates are also economically relevant. The water that is added to milk is sourced from the ground and is thus for free. This means an increase in adulteration of milk due to an increase in the number of over-reported rolls by one standard deviation would yield milkmen higher profits on average by INR 3. Over the course of a month, depending on the size of the dairy and the over-reporting behavior of a milkman, the additional gains correspond to up to 8 percent of a milkman's monthly income.[21]

The strong positive association between dishonesty and adulteration is also observed

[20] 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.

[21] To compute gains from adding water, we calculate additional revenue earned from selling a liter of pure buffalo milk and increasing the level of added water in a liter of milk sold by 3 percentage points. We divide the price charged for one liter of milk by its share of pure milk (1 - added water in percent) and compare this to the quotient of the price and share of pure milk minus the additional 3 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.

in the second sample of milk. Without or with controls, the coefficient on dishonesty is close to 0.77 (s.e. 0.359) and is statistically significant at the 5-percent level.[22] This result implies that a one-standard deviation increase in dishonesty (7.08) increases added water by 5.43 percentage points, which is comparable to the estimate we obtain using the first sample (column 4, Table 4). The two coefficients are not significantly different from each other (p-value = 0.57). These findings point towards an important role of honesty in mitigating adulteration in markets.

IV.B. Robustness Checks

We further assuage the concern that the above results are being driven by other factors by conducting a variety of robustness checks.

Economic variables.– It is plausible that poor economic situation drives milkmen to behave dishonestly in the die game and also to adulterate milk in order to earn higher income. We try and proxy for the economic situation of milkmen in several ways. Our main specification already controls for buffalo herd size, so we start by introducing one by one other proxies like an indicator for agricultural land holding, an indicator for car ownership, an indicator for whether a household member works outside the milk market, and family size.[23] The summary statistics on these variables is reported in Panel D of Table 2.

Table A.5 reports the results. In column 1, we control for agricultural land holding, in column 2 for car ownership, in column 3 for outside option, and in column 4 for family size. The magnitude and significance of the coefficient on dishonesty remains unchanged, whereas the coefficients on economic proxies turn out to be small and statistically insignificant. We are concerned that individually these variables may not proxy well for economic situation, so we introduce these jointly in column 5. This does not lead to any changes in the coefficient on dishonesty. As before, proxies of economic situation remain individually and jointly statistically insignificant. There could be a concern that these proxies may cancel each other out, so in column 6 we introduce their underlying first principal component.[24] As before, the coefficient on dishonesty is robust to the inclusion of the principal component, which itself is statistically insignificant.

Altruism.– Hanna and Wang (2017) and Barfort et al. (2019) find a positive correlation between honesty and altruism. However, these studies are in the context of selection in

[22]Because of the small sample in the second round, we consider a limited set of controls including years in operation and price. The results hold in magnitude and significance even when we consider religiosity, caste, and lactation.

[23]We use an indicator for agricultural land holding because many milkmen were reluctant to provide data on actual land holding. It is very likely that many misreported at the intensive margin but not at the extensive margin, so we prefer using an indicator for this variable.

[24]The principal component is highly and significantly correlated with all four proxies of economic situation.

public service, where altruism, as in motivation to serve the society, can be thought of as playing a role. In our context, milkmen are not expected to enter the milk market because of altruism, as in selling milk for free or at a lower price to the poor and needy. We did not come across any such activity. Also, when milkmen sell quality that is commensurate with price they are simply meeting contractual obligation and are not being altruistic (see Khalil 2003). Nonetheless, we control for other such potentially correlated behaviors that we constructed using dictator games. This includes an indicator for altruism, spite, reciprocity, and others, whereby selfish types serve as a benchmark category. Table A.6 reports the results. Since the experiment was conducted with a sub-set of milkmen, we first verify in column 1 that the association between dishonesty and adulteration is positive and statistically significant in this sub-sample after controlling for the same covariates and fixed effects as in the main specification. In column 2, when we introduce indicators of behavioral types, the magnitude and significance of the coefficient on dishonesty remain stable. In contrast, the coefficients on behavioral types are individually and jointly statistically insignificant (p-value=0.91).

Other control variables.– We conduct a variety of additional robustness checks in Table A.7. There is a possibility that milkmen differ in the quality of fodder fed to buffaloes, which affects milk output. If dishonest milkmen use poor quality fodder and also add water to offset for the drop in milk output, it could result in omitted variables bias. So, we control for fodder quality using monthly expenditure on fodder per buffalo in column 1. A similar concern arises if milkmen with higher production costs add more water to milk to save on these costs and also behave dishonestly in the die game. Thus far, our specification controls for production costs using two separate variables: feeding costs and buffalo herd size. Since milkmen milk and feed their buffaloes themselves, there are no outside labor costs. In column 2, we control for total production costs, which combine feeding costs as well as the cost of buffaloes. To account for scale effects, we also include a quadratic term of this variable. Fisman and Miguel (2007) show that home environment matters for corruption. Accordingly, we control for an indicator of the state of origin of milkmen in column 3. A milkman might behave dishonestly because he thinks his neighbors are dishonest. We test this by including the number of over-reported rolls of the nearest neighbor as an additional control in column 4. We also test for the role of luck by explicitly controlling for the actual outcome of the die rolls in column 5. Finally, we also use an alternative dependent variable which uses net added water after accounting for water added by the nearest neighbor in column 6.[25] The results clearly show that the coefficient on dishonesty remains remarkably stable in magnitude (around 0.46) and significance (p-value < 0.001). In contrast, the coefficients on additional control variables are statistically insignificant. When we introduce controls for the exact time (in hours)

[25]This variable is captured as: added water by a milkman - added water by the nearest neighbor.

at which the sample was bought and the temperature of milk at the time of testing, it does not change any of our main findings (results available on request). This is expected given that we purchased milk from all milkmen in the month of December.

Alternative measures of dishonesty.– We reproduce our results using alternative measures of dishonesty. We start with the sum of added points, which allows us to go beyond incidence and take the magnitude of over-reporting into consideration. Column 1 in Table 5 reports the result and shows that the coefficient is 0.174, which is significant at the 1-percent level. According to this estimate, one standard deviation increase in the sum of added points (15.85) leads to an increase in added water by 2.76 percentage points.

Our measures of dishonesty could be partly reflecting 'bad luck' in the experiment. To mitigate this, we construct two additional measures, which express the number of over-reported rolls and the sum of added points in relative terms. We construct the variable share of over-reported rolls by dividing the number of over-reported rolls by the number of recorded rolls in which an individual did not obtain a '6', as in this case over reporting is not possible. 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. Columns 1-2 of Table A.8 report the results and show that the association of these two measures of dishonesty with added water in milk remains positive and highly statistically significant (p-value < 0.001). Notably, standardized coefficients show that estimates obtained from different measures are comparable to each other and fall between 2.6 and 3.1. Notice that the coefficient on price always remains statistically insignificant.

Bluetooth misses.– Due to temporary connectivity deficiencies, the Bluetooth die did not transmit the actual outcome for 10 percent of the die rolls. Since technology shocks are expected to be random, we doubt if they have a bearing on our results. In any case, our continuous measures of dishonesty account for the potential concerns that arise from missed rolls. Still, we conduct a battery of robustness checks to show that this is indeed the case. Table A.9 reports the results. The coefficient on the number of over-reported rolls is close to 0.46 and remains significant at the 1-percent level. All of these results also hold when we use alternative measures of dishonesty.

IV.C. Dishonesty, Adulteration, and Nutrition

Though it is beyond the scope of this study to show the negative consequences of dishonesty on nutrition via the addition of water to milk. Nonetheless, in Table A.10 we highlight its negative association with the amount of protein (column 1) and micronutrients (SNF) in milk (column 2). A one-standard deviation increase in the number of over-reported rolls (6.6) leads to a fall in protein content of milk by 0.20 percentage points and in SNF by 0.43 percentage points. Given that the average levels of these nutrients in buffalo milk are 4.6 and 9.5 percent, these losses are not trivial.

We substantiate this point further for the protein loss using back of the envelope calculations. According to the Indian Council of Medical Research, the daily recommended dietary allowance for Indians is 80 grams of protein of which 14 grams is expected to come from milk (Rao 2013). However, if the milk is diluted such that it contains 18 percent water, then an individual actually obtains 11 grams of protein. This leads to a deficit of 3 grams or nearly 20 percent of the recommended intake of protein from milk.

IV.D. Is the Bluetooth-based information warranted?

The Bluetooth die allows us to construct measures of dishonesty that bypass the measurement error stemming from the randomness of the device. However, the Bluetooth die imposes limits on sample size, as we need to record the actual outcomes of die rolls for each participant.[26] A normal die offers the possibility to obtain a larger sample, but it allows us to construct measures of dishonesty that suffer from stronger measurement error. It might be possible to use an instrumental variables approach to address this problem, but this depends on whether the instrument is valid. Therefore the advantage of using a Bluetooth die depend on the size of the measurement error in self-reported measures and the possibilities to address this problem.

To shed light on these topics, we construct a self-reported measure used in previous studies. It is based on self-reported points that a milkmen accumulates over the 40 rolls. According to this measure, the larger the number of self-reported points the more dishonest a milkman is expected to be. The average number of self-reported points in our sample is 147.47 (s.d. 21.50) which is significantly higher than the expected value of the rolls (140 points). To gauge the extent of measurement error in this variable we plot in Figure 5 self-reported points (y-axis) against a Bluetooth based measure of dishonesty (Sum of over-reported rolls).[27] From the figure, the measurement error appears to be large: many honest milkmen (zero over-reporting) have self-reported points that exceed the expected value (140 points), whereas many dishonest milkmen have self-reported points that fall below the expected value.

[26]It may be possible to overcome this handicap by using several Bluetooth dice, but one would now have to additionally ensure a minimum distance between participants so that the Bluetooth signals do not interfere with each other.

[27]These two measures are comparable as they measure dishonesty at the intensive margin.

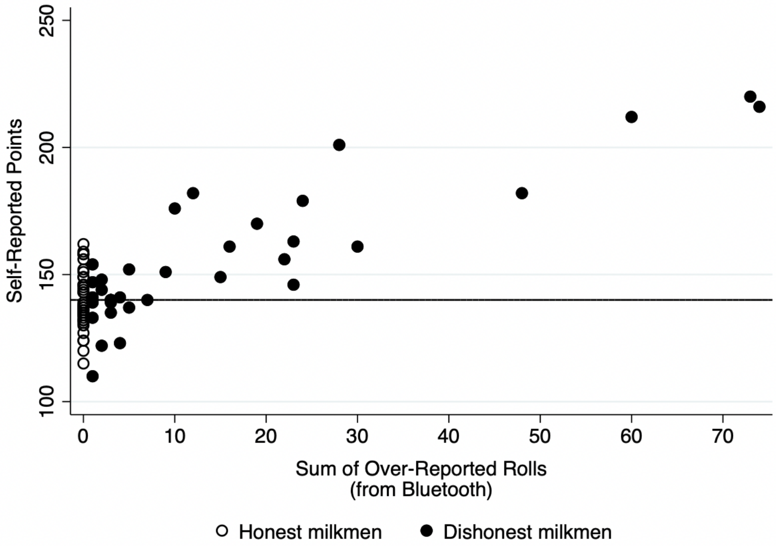


Figure 5: Self-Reported Points and Sum of Over Reported Rolls

Notes. The solid horizontal line indicates the expected value of the outcome over 40 die rolls.

Also, notice that though the two measures are positively correlated, it is not 1, thereby suggesting that the self-reported measure is prone to measurement error. Indeed, when we regress the sum of over-reported rolls on self-reported points, the coefficient on the latter (reliability ratio) turns out to be 0.61, suggesting that the measurement error is large (see Angrist and Kruger 1999). These results suggest proportional attenuation bias (1-reliability ratio) of 39 percent.[28] We test for this in column 2 of Table 5, which presents results on the association of self-reported points with adulteration after controlling for covariates and fixed effects. The coefficient on self-reported points is 0.11 and it is statistically significant at the 5-percent level. This is smaller than the coefficient of 0.17 on the sum of over-reported rolls from Bluetooth die (see column 1), suggesting an attenuation bias of roughly 35 percent.

It might be possible to resolve the measurement error problem by using an instrument for self-reported measures under the assumption that the measurement error is classical. Since each milkman rolled a die 40 times, it is possible to use a measure based on second 20 rolls (r_2) as an instrument for a measure based on the first 20 rolls (r_1). r_2 serves as a valid instrument provided the measurement error in r_1 is uncorrelated with r_2. Figure A.2) shows that the association between r_2 and the measurement error in r_1 is positive. A regression analysis shows that the magnitude of the coefficient on the measurement error is 0.39 (s.e. 0.155), which is not only large but also statistically significant (p-value $= 0.014$). This result clearly shows that r_2 is not a valid instrument for r_1 (see Angrist and Kruger

[28]This result is not affected by Bluetooth misses. The measure self-reported points is uncorrelated with the number of Bluetooth misses. Also, the results remain unchanged when we construct reliability ratio using only non-missed Bluetooth rolls.

1999).[29] This means the instrumental variables approach may not be particularly useful in our context.[30] Together, these results suggest that the Bluetooth-based information is warranted as it is useful in mitigating the measurement error from the randomness of the device and resulting attenuation bias in self-reported measures of dishonesty.

Even though the IV approach is not appropriate to correct for the measurement error in self-reported outcomes, it may be useful to correct for measurement error in Bluetooth-based outcomes from the randomness in behavior. We construct an IV analogous to the approach described in the above paragraph. In the regression of added water on sum of over-reported rolls, the IV coefficient turns out to be 0.193 (s.e. 0.072) and it is statistically significant at the 1-percent level. This result indicates the the attenuation bias from the randomness in behavior is small.

V. Verifiability of Milk Quality

The role of honesty in mitigating adulteration gains prominence if milk quality sold in markets is not easily verifiable. We investigate this via incentivized milk-testing tournaments that we conducted separately with milkmen and customers. The professional laboratory testing of milk is costly therefore individuals are likely to verify milk quality based on simple tests which can be implemented at home. In the milk-testing tournaments, milkmen and customers could taste the milk and/ or test its viscosity by letting milk flow down their palm and fingertips.

In the tournament, milkmen and customers were presented five different samples containing 0, 100, 200, 300, and 400 ml of clean water added to a liter of pure buffalo milk. Barring the first four samples, the fifth sample is at the threshold beyond which adulteration is perceptible. These samples also reflect the range in added water in the milk samples we purchased from milkmen.

We asked individuals to come one by one to a room and predict the amount of water added in each of these five samples or rank these on quality (1 being the best and 5 being the worst). In each tournament, the three individuals whose predictions were closest to the actual level of added water were paid INR 800, INR 500, and INR 300 respectively. These earnings (worth 14, 9 and 5 liters of milk sold or consumed) were high enough for both milkmen and customers to take the tournament seriously. This also seems to be the case, as each individual spent on average 4-5 minutes in figuring out the milk quality. Moreover, individuals gained prestige from winning the tournament. In each

[29] If we ignore this concern and proceed with the IV analysis, the coefficient on self-reported points turns out to be only 0.124, which though statistically significant is only slightly larger than its OLS counterpart (see column 1 of Table 5) but much smaller than the coefficient on Bluetooth-based measure (see column 2 of Table 5).

[30] As before, this result is not due to Bluetooth misses and remains unchanged when we take Bluetooth misses into account to reconstruct our variables.

neighborhood, fellow milkmen lifted the winner in the air; among the customers, the winners had to invite fellow neighbors for a treat.

In two neighborhoods, milkmen predicted the added water in milk (44 milkmen). In the third neighborhood milkmen only ranked the milk quality but did not predict the level of added water (24 milkmen). In the fourth neighborhood, we conducted a modified tournament in which milkmen were presented two samples containing 250 ml, another two containing 350 ml, and the last sample containing 400 ml of added water (30 milkmen). This modification allows us to test the accuracy of predictions in samples containing the same level of added water.

V.A. Empirical specification

To test the verifiability of milk quality, we use the following OLS specification:

$$P_{is} = \delta_0 + \delta_1 Actual_s + \delta_i + \mu_{is} \tag{2}$$

where p_{is} is the predicted value (predicted rank) by individual i of milk sample s. Actual is the true level of added water (rank) in the sample. δ_i is the fixed effect for the individual, and μ_i is an error term. The coefficient of interest is δ_1, which measures the association between predicted and actual levels (rank) of added water. To the extent, milkmen and customers can verify milk quality, then δ_1 is expected to be positive. If customers and milkmen could fully observe the actual quality then this coefficient will be equal to 1.

A key concern in estimating the above equation is that if individuals do not take the tournament seriously then this could lead to an erroneous interpretation that milk quality is not verifiable. We suspect such concerns to arise if milkmen who sell poor quality want to deceive us purposefully by reporting arbitrary values of added water or if wealthier individuals give noisy estimates because the stakes are not high enough for them. We doubt that these concerns hold water because both customers and milkmen were very keen on taking part in the tournament and spent several minutes figuring out the milk quality. Also, in addition to earning money individuals acquired prestige from winning the tournament. Nevertheless, because each individual took five decisions, we can include individual fixed effects to account for such concerns. Thus, to the extent δ_i is included in the specification, δ_1 can be considered as yielding an unbiased estimate of the verifiability of milk quality.

V.B. Results

We presented results from the milk testing tournament first using predicted level and then predicted rank. We also conduct heterogeneous analysis to investigate if the results differ by dishonesty and milk flavor.

Predicted levels.– The upper panel of Figure 6 shows the average milk quality predicted by customers and milkmen for each level of actual milk quality. Two findings are noteworthy. First, although customers and milkmen differ in their average prediction, it reflects a constant differential. Second, the average predicted level seems to be more or less flat for both customers and milkmen. This suggests that predicting the level of added water in milk is indeed difficult.

We test these patterns econometrically in Table 6. Column 1 considers only the sample of customers and shows that the coefficient on actual level is very small in magnitude and is also statistically insignificant. Column 2 considers only the sample of milkmen who predicted quality and shows that the coefficient on actual level is also small in magnitude and statistically insignificant. Moreover, the null hypothesis that the coefficient on actual level is not significantly different between customers and milkmen cannot be rejected (p-value = 0.27). Since the coefficients do not change when we introduce individual fixed effects separately for milkmen and customers, we combine the sample in column 3 and introduce individual fixed effects. The coefficient on actual level remains small and statistically insignificant.

The results do not change when we use average absolute deviation of predicted water from the actual water in milk (see Figure A.3). The absolute deviation for both customers and milkmen turns out to be large, and there is no overall difference in their patterns (p-value = 0.37). We also look at share of individuals whose predictions fall within $\pm$ 20 percent of the actual level. Except for one customer and one milkman, the deviations exceeds 20 percent for all customers and milkmen (average deviation is 70 percent).[31]

[31] It is not possible to construct deviation in percent when the actual level of added water is zero, so we exclude these cases. If we were to consider this sample as well then even the two individuals whose deviations are within 20 percent are excluded because they predicted pure milk as having 50 ml and 150 ml of added water.

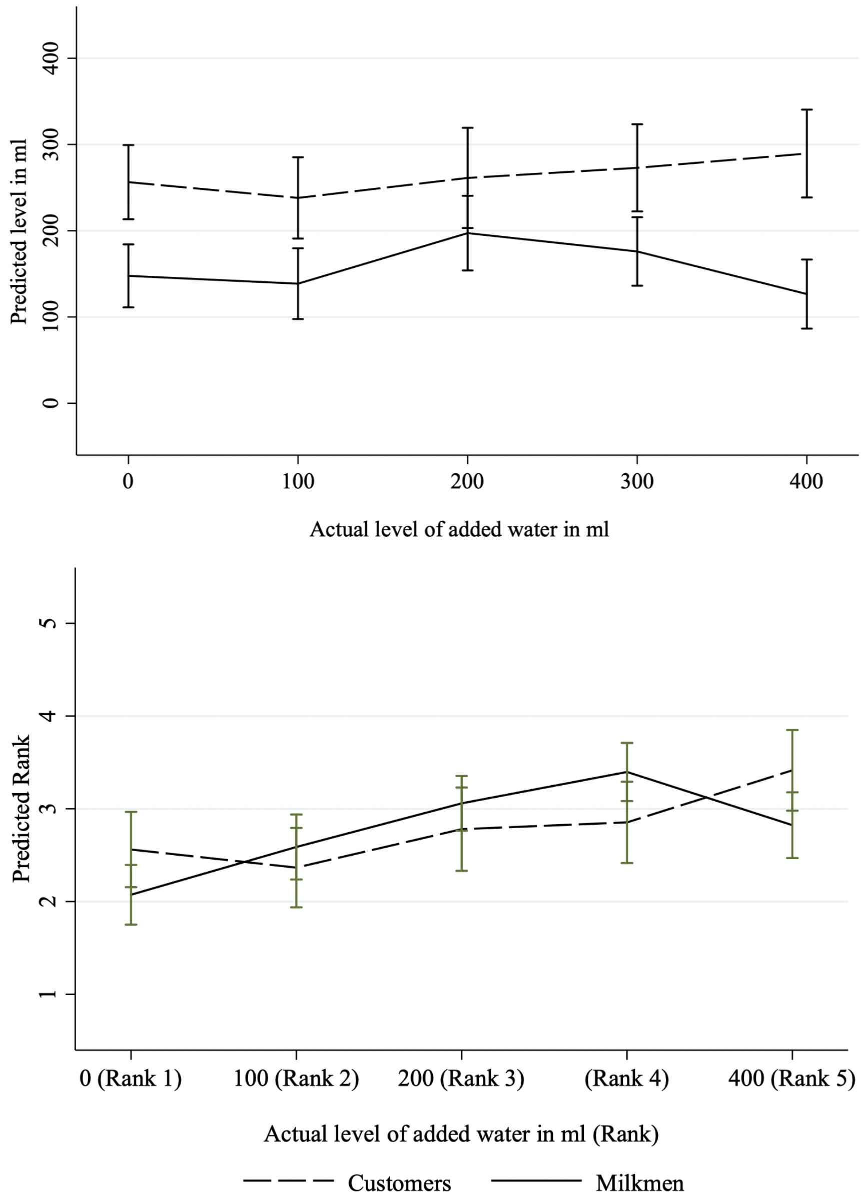


Figure 6: Predicted levels of added water and ranking of milk quality in the milk-testing tournament

Notes. Y-axis presents the predicted level (rank) of added water in a liter of milk. X-axis presents the actual level of added water to a liter of milk by us.

Similar results are obtained when we consider the alternative tournament in which we presented milkmen with different samples containing the same level of added water. Figure A.4 shows that milkmen predict the same quality as having different levels of added water. While the predictions are not significantly different for samples containing 250 ml of added water, they are significantly different for samples containing 350 ml of added water. When we regress predicted level on actual level using the alternative sample, the coefficient on actual turns out to be small and statistically insignificant (coef. 0.116 s.e. 0.093).

Predicted ranks.– It could be that while milkmen and customers are not able to predict precise levels of added water, they are able to rank milk quality. So, we investigate whether there exists an association between predicted and actual rank. For this analysis, we transform data from predicted levels to ranks (as necessary).

We plot in the lower panel of Figure 6 the relationship between predicted and actual rank separately for milkmen and customers, where 1 is the highest rank and 5 the lowest. There is a positive association between predicted and actual rank, suggesting that poor quality milk samples are ranked lower than higher quality ones. The Spearman rank correlation confirms this for both customers (ρ = 0.22, p-value$<$0.01) and milkmen (ρ = 0.23, p-value$<$0.01), but the magnitude of the correlation is rather small. When we examine these patterns parametrically with or without individual fixed effects in Table 6 they yield similar results.

The overall patterns mask important differences in how customers and milkmen rank milk quality. These differences are relevant for understanding adulteration of milk by milkmen. Using a Chi square test, we find that milkmen rank samples containing over 200 ml of added water significantly lower than the pure milk sample (p-value$<$0.01). In contrast, customers rank only the most diluted sample (400 ml of added water) significantly lower than the pure milk sample (p-value = 0.023). These results imply that customers can distinguish between the ranks only if the milk quality is beyond the threshold at which adulteration is perceptible, but not when it is below the threshold. This also explains why we rarely observe milkmen adding water beyond 30 percent. [32]

Together, these findings suggest that it is difficult to verify milk quality especially if milkmen add water below the threshold.

Heterogeneous analysis.– We carry out heterogeneous analysis along two dimensions. First, our results equate water percentage with quality, but there could be other differences that consumers care about. One such difference could be flavor. To the extent flavor does not correlate perfectly with quality, this could affect the interpretation of our results. Second, although our results control for individual specific differences via individual fixed effects, it might be interesting to see if prediction patterns vary by the dishonesty of milkmen.

From the surveys we conducted with milkmen and customers, quality and flavor appear to be correlated (see Section II). Fortunately, we can also test this empirically because we noted for each participant whether they tasted milk or not. 55 percent of the milkmen who took part in the tournament tasted milk; the corresponding share for customers is 49 percent. Figure A.5 shows that there is no difference in the results either for milkmen (top panel) or for customers (lower panel). In a regression of predicted level (rank) on

[32] Of the 172 milk samples that we collected (72 and 29 samples from round 1 and 2 of the main study and 71 samples from the pilot study) only 7 (4 percent) had added water above this threshold.

actual level (rank) and an indicator for tasted milk, the interaction term turns out to be statistically insignificant (p-value = 0.56).[33]

We now turn to analyzing the behavior of honest and dishonest milkmen in the milk-testing tournament. To leverage a larger sample, we also include milkmen from the neighborhood where the alternative tournament was conducted.[34] We classify milkmen as dishonest if the number of over-reported rolls exceeds zero, otherwise honest. Figure A.6 shows that there are no differences by honest and dishonest milkmen in the milk-testing tournament with respect to both predicted levels and ranks.

VI. Conclusions

In markets with asymmetric information and poor enforcement of rules, sellers have a strong incentive to cheat by selling poor quality at the price of good quality. This is a serious concern, particularly when the market is for food products and public health is at stake. Despite this incentive, there is large and persistence variation in the product quality sold by sellers. This paper examines the role of innate honesty in explaining this variation using the context of buffalo milk markets in India.

Milk is a credence good, which creates strong incentives for milkmen to adulterate milk with water and sell it at the price of pure milk. We assess the extent of adulteration in milk sold in markets using an ultrasonic milk analyzer. We find large variation in added water, which reflects purposeful adulteration. We proceed by measuring dishonesty via a behavioral experiment in which milkmen have to roll a die and then self-report its outcome. We use a Bluetooth enabled die, which allows us to obtain actual outcomes of the die rolls. We contrast self-reported and actual outcomes to obtain our measures of dishonesty. This approach allows us to mitigate the concern over errors in variables from the randomness of the device that the self-reported measures are prone to.

We find a strong positive and significant association between dishonesty and added water in milk, which is robust to the inclusion of a powerful set of controls. Similar results are obtained when we use the second sample of milk collected for a sub-set of milkmen. Our findings highlight the importance of honesty in situations where monitoring and reputation have a limited role to play. It shows that some sellers are willing to forego private monetary gains in return for utility they get from being honest in an environment that encourages cheating.

We then test whether milk quality is verifiable through an incentivized milk-testing tournament. We present milkmen and customers with samples of milk containing different

[33]Similar results are obtained when we analyze data from the neighborhood in which we conducted the alternative tournament (results available on request).

[34]The data on behavior in the die game is not available for all milkmen who took part in the tournament. This is because the two samples overlap but are not exactly the same. We could match the behavior of 35 milkmen. Of these, 19 are honest and 16 are dishonest.

levels of added water and then ask them to predict the same. Our results reveal that both milkmen and their customers are unable to predict the actual level of added water in milk or the rank of the milk quality that is typically sold in markets. However, we do find that the worst quality that is off the equilibrium and hence not sold in markets is ranked slightly lower than the best quality.

Under the existing set-up, price collusion prevents honest milkmen from signaling quality via price. This in itself does not drive out honest milkmen from the market, but it has implications for welfare and public policy. While there exist numerous institutions to counteract the effect of uncertainty in product quality including limited liability, guarantees, and branding, they may have a limited role under the current set up. Future studies could guide public policy by investigating consumer side and milkmen side interventions that benefit honest milkmen and customers. Potential intervention could include installation of machines in areas where the customers reside. While implementing this policy it is important to note that milk quality is multidimensional. This means that if only added water in milk is tested then milkmen may switch to other adulterants that are currently not being used and cannot be detected by the machine. Therefore, one could combine this with periodic certification of milk quality by third parties (see Dranove 2012, Duflo et al. 2013). It is unlikely that milkmen will be able to game the disclosure system like hospitals (denying severally ill patients). Such a policy may help break collusive arrangements, allow honest milkmen to charge a higher price, and let customers switch according to their preference.

References

Angrist, J.D. and A. B. Krueger. 1999."Empirical Strategies in Labor Economics." in Handbook of Labor Economics by O. AshenJelter and D. Card. Elsevier, 1277-1366.

Arrow, K. 1963. "Agency and the welfare economics of medical care." *American Economic Review*, 53, 941-973.

Athey, S. and Bagwell, K. 2001. "Optimal Collusion with Private Information." *Rand Journal of Economics*, 32, 428-465.

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

Andreoni, J. 1988. "Why Free Ride? Strategies and Learning in Public Goods Experiments." *Journal of Public Economics*, 37, 291–304.

Baker, G., R. Gibbons, and K. J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics*, 109, 1125–56.

Banerjee. A. and E. Duflo. 2000. "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry." *Quarterly Journal of Economics*, 115, 989-1017.

Banerjee, A., Hanna, R., and Mullainathan, S. 2012. "Corruption." Handbook of Organizational Economics. Ed. Robert Gibbons and John Roberts. Princeton University Press, 1109-1147.

Barfort, S., Harmon, N.A., Hjorth, F., and Olsen, A.L. 2019. "Sustaining Honesty in Public Service: The Role of Selection." *American Economic Journal: Economic Policy*, 11, 96–123.

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

Bertrand, M. and Mullainathan, S. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94, 991-1013.

Bloom, N. and J.V. Reenen. 2007. "Measuring and Explaining Managament Practices across Firms and Countries." *Quarterly Journal of Economics*, 122, 1351-1408.

Buettel, B., M. Fuchs, B. Holz. 2008. "Freezing Point Osmometry of Milk to Determine the Additional Water Content – an Issue in General Quality Control and German Food Regulation." *Chemistry Central Journal*, 2, 1-7.

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

Cason, T.N. and S. Y. Wu. 2019. "Subject Pools and Deception in Agricultural and Resource Economics Experiments." *Environmental and Resource Economics*, 73, 743-758.

Charness, G. and M. Rabin. 2002. “Understanding Social Preferences with Simple Tests.” *Quarterly Journal of Economics*, 117, 817-869.

Chen, Y. and S.X. Li. 2009. “Group Identity and Social Preferences.” *American Economic Review*, 99, 431–457.

Cohn, A. and M. Marechal. 2018. “Laboratory Measure of Cheating Predicts Misbehavior at School.” *Economic Journal*, 128, 2743-2754.

Dai, Z., Galeotti, F., and Villeval, M.C. 2017. “Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportations.” *Management Science*, 64, 1081-1100.

Das, J., Holla, A., Mohpal, A., Muralidharan, K. 2016. “Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India.” *American Economic Review*, 106, 3765-99.

DDA. 1985. Mini Master Plan: “Integrated Development of Urban and Rural Villages in Delhi.” Delhi Development Authority.

DDA. 2006. “Appeal No.CIC/WB/A/2006/00317.” Right to Information Act, Central Information Commission.

DDA. 2018. “DDA notifies charges for allowing permitted land use change in Delhi’s industrial clusters.” http://delhi-masterplan.com.

Dizon-Ross, R., Dupas, P., and Robinson, J. 2015. “Governance and Effectiveness of Public Health Subsidies.” Working Paper 510, Stanford Center for International Development.

Dranove, D. 2012. “Health Care Markets, Regulators, and Certifiers.” Handbook of Health Economics. Ed. Mark Pauly, Thomas McGuire, and Pedro Barros. Elsevier, 639-690.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. 2013. “Truth-Telling by Third Party Auditors and the Response of Polluting Firms: Experimental Evidence from India.” *Quarterly Journal of Economics*, 1499-1545.

Dulleck, U. and Kerschbamer, R. 2006. “On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods.” *Journal of Economic Literature*, XLIV, 5-42.

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

Fafchamps, M. 2004. “Market Institutions and Sub-Saharan Africa: Theory and Evidence.” Cambridge: MIT Press.

Fafchamps, M. 2010. “Spontaneous Markets, Networks and Social Capital: Lessons from Africa” in Institutional Microeconomics of Development, Timothy Besley and Rajshri Jayaraman (eds.), CESifo Seminar Series, Cambridge: MIT Press.

Faraz, A., Lateef, M., Mustafa, M.I., Akhtar, P., Yaqoob, M., and Rehman S. 2013. "Detection of adulteration, chemical composition and hygienic status of milk supplied to various canteens of educational institutes and public places in Faisalabad." *Journal of Animal Plant Science*, 23, 119–124.

Fehr, E., Gächter, S., and Kirchsteiger, G. 1997. "Reciprocity as a Contract Enforcement Device: Experimental Evidence." *Econometrica*, 65, 833-860.

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

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

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

Gächter, S. and J. Schulz. 2015. "Intrinsic Honesty and the Prevalence of Rule Violators across Societies." *Nature*, 531, 496 – 499.

Glennerster, R. and Powers, D. 2014. "Balancing Risk and Benefit: Ethical Tradeoffs in Running Randomized Evaluations" in Handbook of Professional Economic Ethics, edited by George DeMartino, Oxford University Press.

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

Hanna, R. and Wang, S.-Y. 2017. "Dishonesty and Selection into Public Service." *American Economic Journal: Policy*, 9, 262-290.

Harmon, N.A. 2021."Measuring Individual Dishonesty Using (Repeated) Probabilistic Lying Games." Working Paper.

Karthik, N. 2009. "Strategic Communication with Lying Costs." *Review of Economic Studies*, 76, 1359 – 1395.

Liu, J., Ren, A., Yang, L., Gao, J., Pei, L., Ye, R., Qu, Q., and Zheng, X. 2010. "Urinary tract abnormalities in Chinese rural children who consumed melamine-contaminated dairy products: a population-based screening and follow-up study." *Canadian Medical Association Journal 182*, 439–42.

Macchiavello, R. and Morjaria, A. 2015. "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports." *American Economic Review 105*, 2911-2945.

Ménard, O., Ahmad, S., Rousseau, F., Briard-Bion, V., Gaucheron, F., Lopez, C. 2010. "Buffalo vs. cow milk fat globules: Size distribution, zeta-potential, compositions in total fatty acids and in polar lipids from the milk fat globule membrane." *Food Chemistry*, 120, 544-551.

Munshi, K. 2007. "From Farming to International Business: The Social Auspices of

Entrepreneurship in a Growing Economy." NBER Working Paper 13065.

Nagin, D. S., Rebitzer, J. B., Sanders, S., and Taylor, L.J. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review*, 92, 850-873.

McMillan, J. and Woodruff, C. 1999. "Interfirm Relationships and Informal Credit in Vietnam." *Quarterly Journal of Economics*, 114, 1285 – 1320.

Olken, B. 2007. "Monitoring Corruption: Evidence from a field Experiment in Indonesia." *Journal of Political Economy* 90, 853-870.

Oster, E. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 37, 187-204.

Platteau, J.P. 2000. "Institutions, Social Norms, and Economic Development." London: Hardwood Academic Publishers.

Rao, B.S.N. 2013. "Fruits, Vegetables, Milk, and Animal Foods in Balanced Indian Diets – A Critical Appraisal." *Bulletin of the Nutritional Foundation of India*, 34, 1-8.

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, 181-196.

Satterthwaite, M. 1979. "Consumer information, equilibrium industry price, and the number of sellers." *Bell Journal of Economics*, 10, 483-502.

Sheldon, I.M. 2017. "Certification Mechanisms for Credence Attributes of Foods: Does It Matter Who Provides Diagnosis,?" *Annual Review of Resource Economics*, 9, 33–51.

Staal, S.J., A.N. Pratt, and M. Jabbar. 2008. Dairy Development for the Resource Poor. A Comparison of Dairy Policies and Development in South Asia and East Africa. Pro-poor Livestock Policy Initiative (PPLPI), Working Paper No. 44-1.

Souza SS, Cruz AG, Walter EHM, Faria JAF, Celeghini RMS, Ferreira MMC, Granato D, Sant'Ana ADS. 2011. "Monitoring the authenticity of Brazilian UHT milk: A Chemometric Approach." *Food Chemistry* 124, 692–5.

Table 2: Summary Statistics

	Mean (1)	Standard Deviation (2)	Coef. on Dishonesty (3)	Standard Error (4)
	A: Adulteration			
Added water in milk (%)	17.962	7.488	0.291	0.131
	B: Dishonesty			
Number of over-reported rolls	3.625	6.604	–	–
	C: Main Covariates			
Years in operation	25.597	8.725	0.019	0.158
Price per liter	57.715	4.278	−0.009	0.077
Age	33.847	11.089	−0.140	0.200
Education	0.667	0.475	0.005	0.009
Caste	0.569	0.499	0.008	0.009
Religiosity	21.227	36.702	0.205	0.664
Buffalo herd size	18.778	36.619	−0.156	0.664
Lactation period	0.792	0.409	0.007	0.007
	D: Other Covariates			
Land holding	0.736	0.444	−0.009	0.008
Car Ownership	0.333	0.475	0.000	0.009
Outside option	0.423	0.497	−0.003	0.009
Family size	6.708	2.651	−0.085	0.047
Feeding costs	5.990	3.132	−0.024	0.058
Total production costs	238.685	155.274	3.554	2.825

Notes: Added water in milk is the percentage of water in a liter of buffalo milk purchased from the milkmen. Number of over-reported rolls is the number of rolls in which self-reported outcome exceeds the actual outcome. Years in operation is the number of years a dairy is in operation. Price is the amount paid for a liter of milk. Age is measured in years. Education is an indicator variable, which equals 1 if a milkman has high school education, otherwise 0. Caste is an indicator variable, which equals 1 if a milkman belongs to Gujjar and Jat caste groups, otherwise 0. Religiosity is the number of visits to a temple or a mosque in a month. We do not separately control for religion as we have few Muslims and all of them are from the same neighborhood; we account for this via neighborhood fixed effects. Buffalo herd size is the number of adult buffaloes owned by a milkman. Lactation period is an indicator variable equal to 1 if any of the buffaloes was in the first three months of lactation period in December 2014, otherwise 0. Data on lactation period was not available for two milkmen and was consequently imputed. The mean of this variable without the imputed values is 0.79 (s.d. 0.41). All our results hold, if we drop these two observations. Land holding, car ownership, and outside option are indicators variables. Family size is the numbers of family members living in the same house. Feeding costs are fodder costs per buffalo per month/ 1000 (INR). Total production costs are feeding as well as livestock costs per month/1000 in (INR). Columns 1-2 report the mean and standard deviation of each variable. Columns 3-4 report the coefficient on dishonesty and its standard error obtained from a regression of each covariate on dishonesty.

Table 3: Predicting Dishonesty

	Coefficient on covariates	
	No Fixed Effects	With Fixed Effects
	(1)	(2)
Years in operation	0.037	−0.037
	(0.079)	(0.094)
Price	−0.006	−0.148
	(0.179)	(0.169)
Age	−0.079	−0.064
	(0.063)	(0.065)
Education	0.891	0.895
	(1.283)	(1.409)
Caste	1.600	−1.378
	(1.522)	(1.762)
Religiosity	0.011	0.025
	(0.028)	(0.029)
Buffalo herd size	2.357	−0.018
	(1.606)	(0.012)
Lactation period	2.357	2.834
	(1.606)	(1.804)

Notes: Column 1 reports coefficient on covariates obtained from a regression of dishonesty on covariates that are introduced simultaneously. Column 2 replicates column 1 after additionally controlling for neighborhood fixed effects.

Table 4: Dishonesty and Adulteration in milk Markets

	Added Water in Milk in Percent					
	No controls	Neighbor -hood FE	Controls, assistant FE	Price	Same price per liter	Drop influential obs
	(1)	(2)	(3)	(4)	(5)	(6)
Dishonesty	0.291	0.324	0.490	0.475	0.508	0.471
	(0.158)	(0.150)	(0.145)	(0.142)	(0.200)	(0.141)
				[0.001]	[0.030]	[0.000]
Years			−0.124	−0.114	−0.062	−0.112
			(0.075)	(0.078)	(0.155)	(0.079)
Age			−0.012	0.001	0.025	0.021
			(0.063)	(0.065)	(0.117)	(0.066)
Education			−3.127	−2.768	−2.592	−2.469
			(1.693)	(1.703)	(2.425)	(1.663)
Caste			4.561	4.313	7.612	4.415
			(1.687)	(1.702)	(3.291)	(1.700)
Religiosity			−0.045	−0.038	−0.049	−0.037
			(0.018)	(0.019)	(0.021)	(0.021)
Buffalo herd size			0.025	0.026	0.040	−0.004
			(0.009)	(0.009)	(0.012)	(0.076)
Lactation period			−4.867	−5.000	−6.716	−5.946
			(1.952)	(1.927)	(3.335)	(1.794)
Price				−0.271		−0.163
				(0.203)		(0.174)
Constant	16.905	14.442	24.844	40.054	21.517	34.310
	(0.944)	(1.969)	(3.811)	(12.106)	(4.276)	(10.617)
Neighborhood FE	No	Yes	Yes	Yes	Yes	Yes
Assistant FE	No	No	Yes	Yes	Yes	Yes
R^2	0.07	0.24	0.50	0.52	0.52	0.54
Observations	72	72	72	72	40	68

Notes: OLS regression with robust standard errors in parentheses. Column 2 includes only neighborhood fixed effects, whereas columns 3-6 include both neighborhood and assistant fixed effects. Column 5 considers the sample of milkmen who charge INR 60 per liter of milk and column 6 drops four influential observations. Square brackets below the coefficient on dishonesty in columns 4-6 report p-values from wild bootstrapped clustered standard errors.

Table 5: Comparison of Bluetooth and
Self-Reported Measure of Dishonesty

	Added water in Milk (percent)	
	Sum of over-reported rolls	Sum of self-reported rolls
	(1)	(2)
Sum of over-reported rolls	0.174	
	(0.064)	
Sum of self-reported rolls		0.114
		(0.045)
R^2	0.48	0.46
Controls	Yes	Yes
Fixed effects	Yes	Yes
Obs.	72	72

Notes: OLS regression with robust standard errors in parentheses. Column 1 reports results using the sum of over-reported rolls from the Bluetooth die. Column 2 reports results using the sum self-reported points. Control variables include years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistant.

Table 6: Milk Testing Tournament

	Customers only	Milkmen only	Individual FE
	(1)	(2)	(3)
	Panel A: Predicted Level		
Actual level	0.101	−0.005	0.045
	(0.065)	(0.064)	(0.046)
Constant	243.366	158.170	197.864
	(21.734)	(17.937)	(9.139)
	Panel B: Predicted Rank		
Actual Rank	0.220	0.231	0.221
	(0.082)	(0.063)	(0.050)
Constant	2.137	2.096	2.111
	(0.248)	(0.188)	(0.149)
Individual FE	No	No	Yes
Individuals (Panel A)	41	47	88
Obs. (Panel A)	205	235	440
Individuals (Panel B)	41	68	109
Obs. (Panel B)	205	340	545

Notes: OLS with robust standard errors clustered on the individual. The dependent variable in Panel A is predict level of added water in ml. The dependent variable in Panel B is predict rank of the sample. The sample size for milkmen increases in Panel B because in one neighborhood milkmen had to predict only the rank and not the level. Results do not change when we introduce fixed effects separately for milkmen and customers, so we combine the samples in column 3.

ONLINE APPENDIX:
Measuring Honesty and Explaining Adulteration in Naturally Occurring Markets

Devesh Rustagi and Markus Kroell

Appendix A

I. Procedures for collecting data on adulteration in milk markets

In this section, we provide a description of the procedures that we adopted for collecting data on adulteration in milk markets.

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. This was done because if milkmen add a variety of adulterants to milk then relying exclusively on added water as a measure of adulteration will result in understating the extent of adulteration by some milkmen. We collected milk samples from 105 milkmen from neighborhoods in our study and tested these samples for a broader set of adulterants listed in the FSSAI study (2011). 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. These analyses revealed that water is the only adulterant in milk in Delhi. Moreover, the correlation between estimates of added water by the laboratory and the machine is very strong ($r = 0.93$). As a result, we focus on the percentage of added water in buffalo milk measured using the machine as our field measure of adulteration in milk markets.

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 us precise quantitative measures of added water.

Sampling procedures

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 assistants unknown to the milkmen to execute this task. In each neighborhood, 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 milkman. These bottles were then stored in an icebox to prevent spoilage.

A major concern in collecting these milk samples is locating the dairy farms of the milkmen who took part in our experiment, because most milkmen do not have dairy names or addresses in front of their house. This could result in a mismatch between adulteration in milk markets and dishonesty as well as reputation. In order to avoid this problem, we prepared detailed maps of each neighborhood so that assistants could accurately locate the milkmen we wanted to target. We prepared these maps through guided walks and photographs while conducting the second household survey (Figure A.1). We marked every target milkmen on the map using cues such as color of the house, nearby shops, signs, and pole numbers, etc. For particularly difficult matches, assistants were requested to take pictures of the dairy with their mobile phones, which were subsequently verified using pictures independently obtained by us.

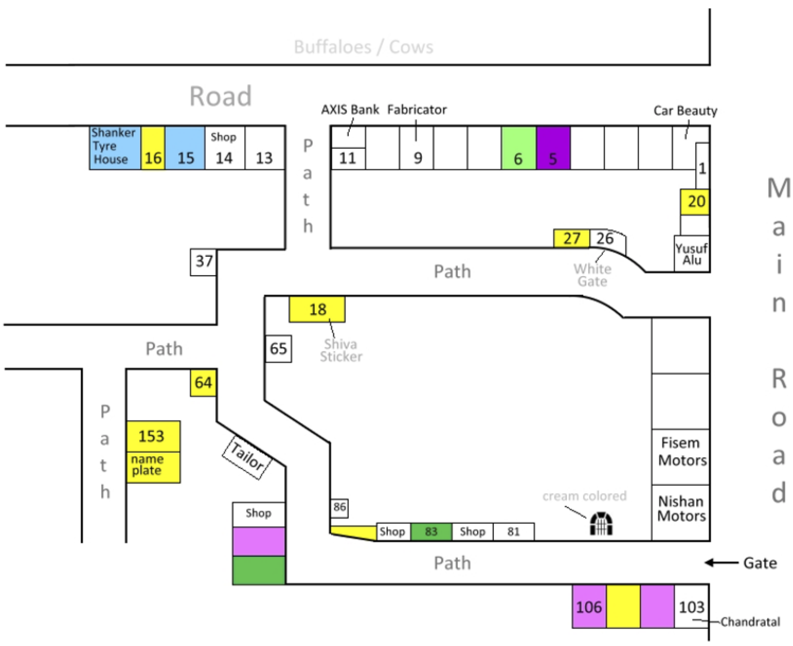


Figure A.1: Map of a neighborhood in our study

II. Sample Construction

Despite our best efforts in collecting milk sample from every milkman who participated in our experiment, we could locate nine milkmen. In addition, we dropped three milkmen for whom data on the number of over-reported rolls is missing due to outages in the Bluetooth-connection. Table A.1 demonstrates that these 12 milkmen do not differ significantly from the 72 milkmen who are in our sample across key socio-demographic characteristics.

Table A.1: Sample Selection - First Milk Sample

	Milkmen in the sample (1)	Milkmen not in the sample (2)	Difference col.(1) - col.(2) (3)
Self-reported points	147.412	145.25	2.222
	(21.504)	(21.872)	(6.721)
Years	25.597	21.093	4.506
	(8.725)	(11.353)	(2.943)
Age	33.847	32.583	1.264
	(11.089)	(9.140)	(3.383)
Education	0.666	0.666	0.000
	(0.475)	(0.492)	(0.148)
Caste	0.569	0.583	−0.104
	(0.499)	(0.515)	(0.156)
Religiosity	21.227	6.917	14.310
	(36.707)	(8.670)	(10.696)
Buffalo herd size	18.778	9.250	9.528
	(35.619)	(12.374)	(10.430)
Lactation period	0.792	0.556	0.236
	(0.409)	(0.527)	(0.149)
Obs.	72	12	84

Notes: Column 1 reports the mean of the covariate for 72 milkmen who are in the sample. Column 2 does the same for 12 milkmen who are not in the sample. Column 3 reports the difference in means between covariates listed in column1 and column 2 using a regression-based test. Columns 1 and 2 report standard deviation in parentheses, whereas column 3 reports standard error in parentheses. For milkmen not in the sample, data on years in operation was not available for 1 milkman and lactation period for 3 milkmen. Price is excluded since it is available for only those milkmen from whom we purchased milk.

We verify in Table A.2 that the 72 milkmen who participated in our study are similar to other small milkmen from their neighborhood along two main characteristics: milk quality and price.

Table A.2: Sample Selection

	Milkmen in the sample (1)	Milkmen not in the sample (2)	Difference col.(1) - col.(2) (3)
Added water	17.962	19.047	2.160
	(7.488)	(8.047)	(1.524)
Price	57.715	58.929	0.297
	(4.278)	(2.721)	(0.694)
Obs.	72	63	135

Notes: Column 1 reports the mean for 72 milkmen who are in the sample. Column 2 does the same for 63 milkmen who are not. Column 3 reports the difference in means using a regression-based test after controlling for neighborhood fixed effects. Columns 1 and 2 report standard deviation in parentheses, whereas column 3 reports standard error in parentheses. Added water is the percentage of water in a liter of milk. Price is the price of a liter of milk.

Since the logistics involved in the collection of milk samples is very demanding, we obtained the second round of sample for a sub-set of milkmen. In Table A.3, we show that the milkmen for whom we have the second sample do not differ from those for whom we do not have the second sample.

Table A.3: Sample Selection - Second Milk Sample

	Coefficient on an indicator for the second sample
Adulteration (1st sample)	−0.847
	(1.906)
Dishonesty (die game)	−1.811
	(1.748)
Years in operation	1.397
	(2.268)
Age	−0.353
	(2.950)
Education	−0.158
	(0.121)
Caste	0.098
	(0.105)
Religiosity	2.517
	(8.906)
Buffalo herd size	−7.826
	(9.268)
Lactation period	0.116
	(0.109)
Price	0.443
	(1.051)

Notes: OLS with standard errors in parenthesis. We compare covariates across 29 milkmen for whom we have the second sample and the remaining 43 milkmen for whom we do not have the second sample. Each row presents results from a separate regression of the covariate listed in the row on an indicator for the second sample as well as neighborhood fixed effects.

III. Robustness of the Main Results

The results in this section show that our results are robust to dropping one neighborhood at a time (Table A.4), controlling for a variety of economic variables (Table A.5), behavioral types including altruists (Table A.6), and other additional controls (Table A.7), using alternative measures of dishonesty (Table A.8), and accounting for Bluetooth misses in a variety of ways (Table A.9).

Table A.4: Adulteration and Dishonesty –
Dropping One Neighborhood at a Time

	Added Water in Milk in Percent					
	Neighbor -hood 1	Neighbor -hood 2	Neighbor -hood 3	Neighbor -hood 4	Neighbor -hood 5	Neighbor -hood 6
	(1)	(2)	(3)	(4)	(5)	(6)
Dishonesty	0.490	0.466	0.480	0.601	0.300	0.487
	(0.164)	(0.156)	(0.184)	(0.159)	(0.128)	(0.143)
Years	−0.083	−0.097	−0.080	−0.121	−0.185	−0.103
	(0.096)	(0.117)	(0.088)	(0.084)	(0.078)	(0.080)
Price	−0.368	−0.389	−0.325	−0.212	−0.149	−0.257
	(0.324)	(0.255)	(0.255)	(0.200)	(0.171)	(0.204)
Constant	48.910	44.667	43.510	38.204	32.952	39.710
	(17.554)	(15.640)	(14.906)	(12.222)	(10.416)	(12.193)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.51	0.55	0.44	0.59	0.55	0.52
Obs.	59	56	60	60	59	66

Notes: OLS regressions with robust standard errors in parentheses. Columns 1-6 report estimates after dropping one neighborhood at a time. Covariates include number of years in operation, price, age, education, caste dummy, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistant.

Table A.5: Adulteration, Dishonesty, and Economic Variables

	Added Water in Milk in Percent					
	(1)	(2)	(3)	(4)	(5)	(6)
Dishonesty	0.481	0.477	0.457	0.485	0.483	0.449
	(0.157)	(0.141)	(0.144)	(0.146)	(0.156)	(0.152)
Land holding	0.353				0.361	
	(1.974)				(2.071)	
Car ownership		0.984			1.470	
		(2.181)			(2.432)	
Outside option			−2.016		−2.562	
			(1.626)		(1.798)	
Family size				0.095	0.120	
				(0.294)	(0.360)	
PC_Economic						−0.273
						(0.765)
Constant	39.636	39.442	40.009	39.504	37.795	40.379
	(12.838)	(12.000)	(12.107)	(12.014)	(12.413)	(12.310)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.52	0.52	0.52	0.52	0.53	0.51
Obs.	72	72	71	72	72	71

Notes: OLS with robust standard errors in parentheses. PC_Economic stands for principal component. The number of observation is less in columns 3 and 6 because we did not have data on the outside option of one milkman. Covariates include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants.

Table A.6: Adulteration, Dishonesty, and Behavioral Types

	Dependent variable: Added Water in Milk	
	Baseline estimates	Behavioral types
	(1)	(2)
Dishonesty	0.489	0.514
	(0.173)	(0.183)
Behavioral types	No	Yes
Controls	Yes	Yes
Fixed effects	Yes	Yes
R^2	0.46	0.47
Obs.	48	48

Notes: OLS with robust standard errors in parentheses. Controls include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants. Behavioral types include an indicator for equality, altruistic, spiteful, and selfish types. Other types is the omitted category.

Table A.7: Adulteration, Dishonesty, and Additional Controls

	Added Water in Milk in Percent					
	feeding costs	Total prod. costs	State of origin	Neighbor's dishonesty	Luck	Alternative dependent variable
	(1)	(2)	(3)	(4)	(5)	(6)
Dishonesty	0.482	0.459	0.468	0.469	0.476	0.535
	(0.154)	(0.155)	(0.146)	(0.147)	(0.142)	(0.197)
Feeding costs	0.040					
	(0.297)					
Total production costs		−0.017				
		(0.023)				
Total production costs squared		0.000				
		(0.000)				
State of origin			−0.778			
			(1.683)			
Neighbor's dishonesty				−0.032		
				(0.154)		
Luck					0.039	
					(0.054)	
Constant	49.553	59.595	41.374	40.255	36.873	33.021
	(28.438)	(25.458)	(13.110)	(12.147)	(13.022)	(17.115)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.53	0.52	0.52	0.52	0.52	0.33
Obs.	63	63	72	72	72	72

Notes: OLS with robust standard errors in parentheses. Controls include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants.

Table A.8: Adulteration and Dishonesty:
Alternative Measures of Dishonesty

	Added Water in Milk in Percent	
	Share of over-reported rolls	Share of added points
	(1)	(2)
Share of over-reported rolls	0.160	
	(0.046)	
Share of added points		0.163
		(0.048)
Years	−0.111	−0.107
	(0.079)	(0.081)
Price	−0.276	−0.268
	(0.209)	(0.213)
Constant	40.111	40.004
	(12.531)	(12.644)
Covariates	Yes	Yes
Fixed effects	Yes	Yes
R^2	0.52	0.51
Obs.	72	72

Notes: OLS with robust standard errors in parentheses. The number of observation is less in columns 3 and 6 because we did not have data on the outside option of one milkman. Covariates include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants.

Table A.9 confirms that Bluetooth misses have no bearing on the interpretation of our results. 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 the number of over-reported rolls is always close to 0.46 and remains highly significant. In column 4, we make the extreme assumption that milkmen always over-reported in all the missed rolls unless they reported a ‘1’. The coefficient on dishonesty remains highly significant at the 1-percent level despite a drop in its magnitude. These results also hold when we use alternative measures of dishonesty.

Table A.9: Adulteration and Dishonesty: Bluetooth Misses

	Added Water in Milk in Percent			
	(1)	(2)	(3)	(4)
Dishonesty	0.464	0.492	0.437	0.308
	(0.152)	(0.154)	(0.166)	(0.112)
Bluetooth miss		−0.136		
		(0.212)		
Covariates	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
r2	0.54	0.52	0.53	0.46
Obs.	72	72	63	72

Notes: OLS with robust standard errors in parentheses. Covariates include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants.

IV. Dishonesty, Adulteration, and Nutrition

Table A.10 shows that protein and SNF (solids not fat) content of milk declines with increase in the degree of dishonesty.

Table A.10: Dishonesty, Protein, and SNF in Milk

	Protein	SNF
	(1)	(2)
Dishonesty	−0.030	−0.065
	(0.009)	(0.019)
Controls	Yes	Yes
Fixed effects	Yes	Yes
Observations	72	72

Notes: OLS with robust standard errors in parentheses. Controls include number of years in operation, price, age, education, caste, religiosity, buffalo herd size, and lactation period. Fixed effects include neighborhood and assistants. SNF stands for solids not fat.

V. IV Validity

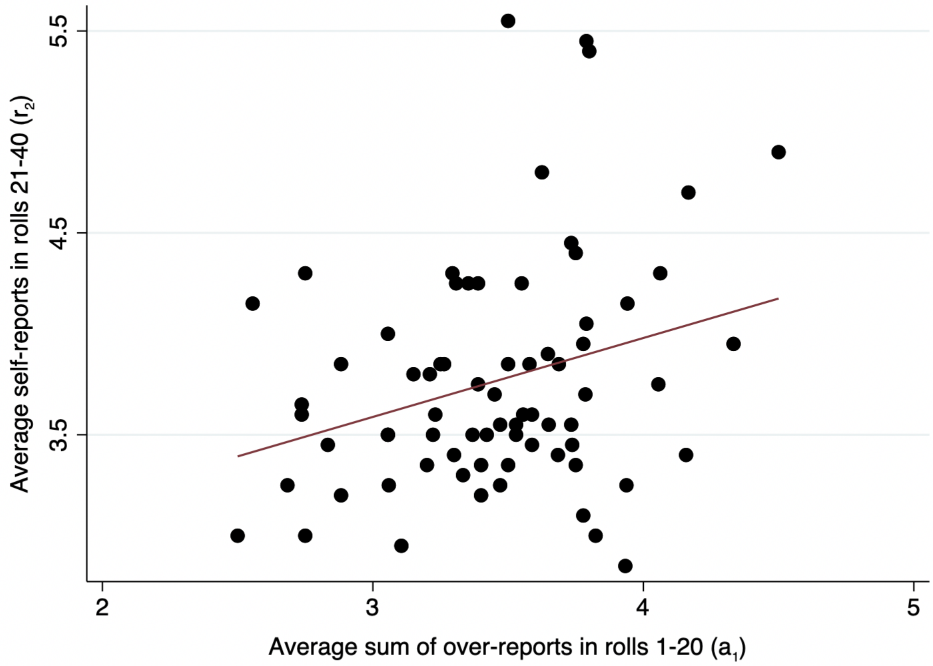


Figure A.2: IV validity (association of r_2 and the measurement error in r_1)

Notes. r_2 is the average of self-reported outcomes in rolls 21-40. The measurement error in r_1 is computed as sum of self-reported outcomes in rolls 1-20 - sum of over-reports in rolls 1-20.

VI. Milk Testing Tournament

Figure A.3 shows the absolute deviation of predicted level from the actual level in the milk-testing tournament. It is evident from the figure that customers and milkmen differ in their deviation only at the widest margin of added water, that is, when the milk samples are pure and highly diluted, but not when the added water takes intermediate values. On average, milkmen have smaller deviation than customers, but the differences are not statistically significant.

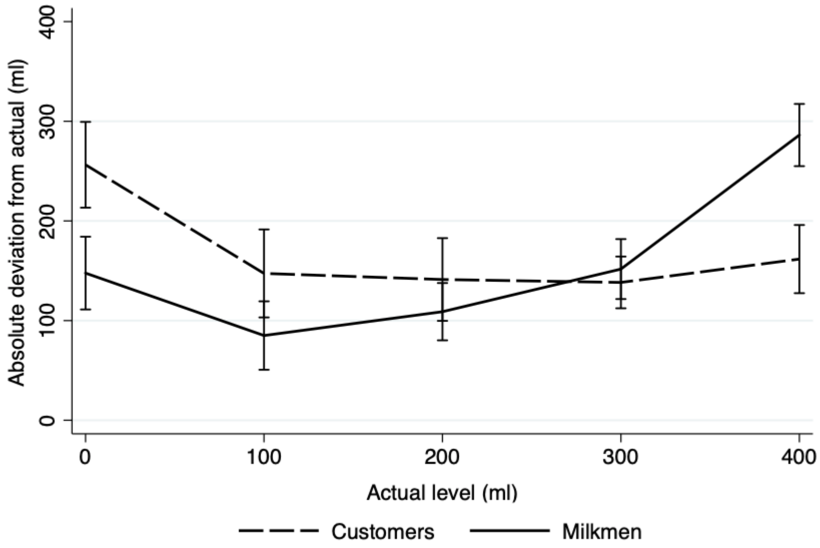


Figure A.3: Absolute deviation of predicted level from the actual level

Figure A.4 shows that milkmen predict samples containing the same level of added water differently, even though the differences are not statistically significant. As with Figure 6 in the main paper, predictions across samples are not statistically significantly different.

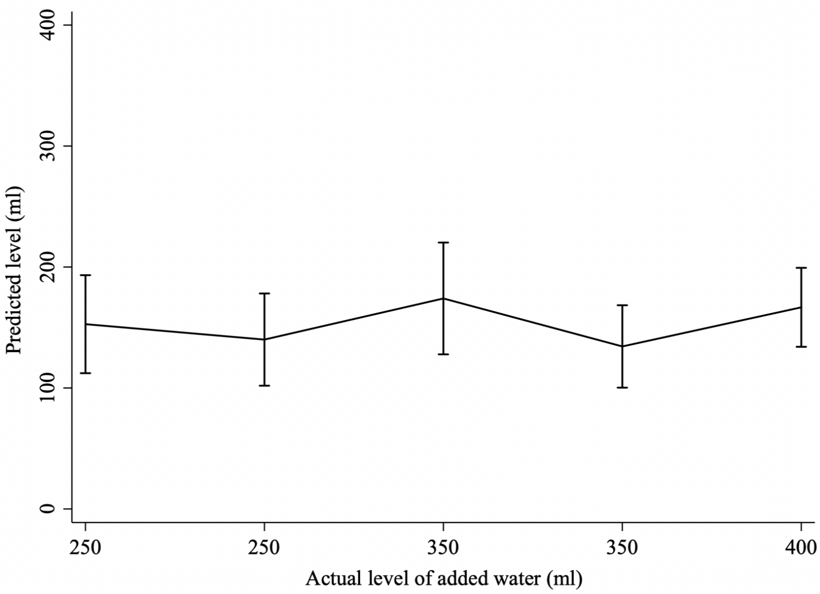


Figure A.4: Predicted levels of added water in a liter of milk

Figure A.5 shows that there is no difference in predicted levels and ranks by milkmen and customers who tasted milk.

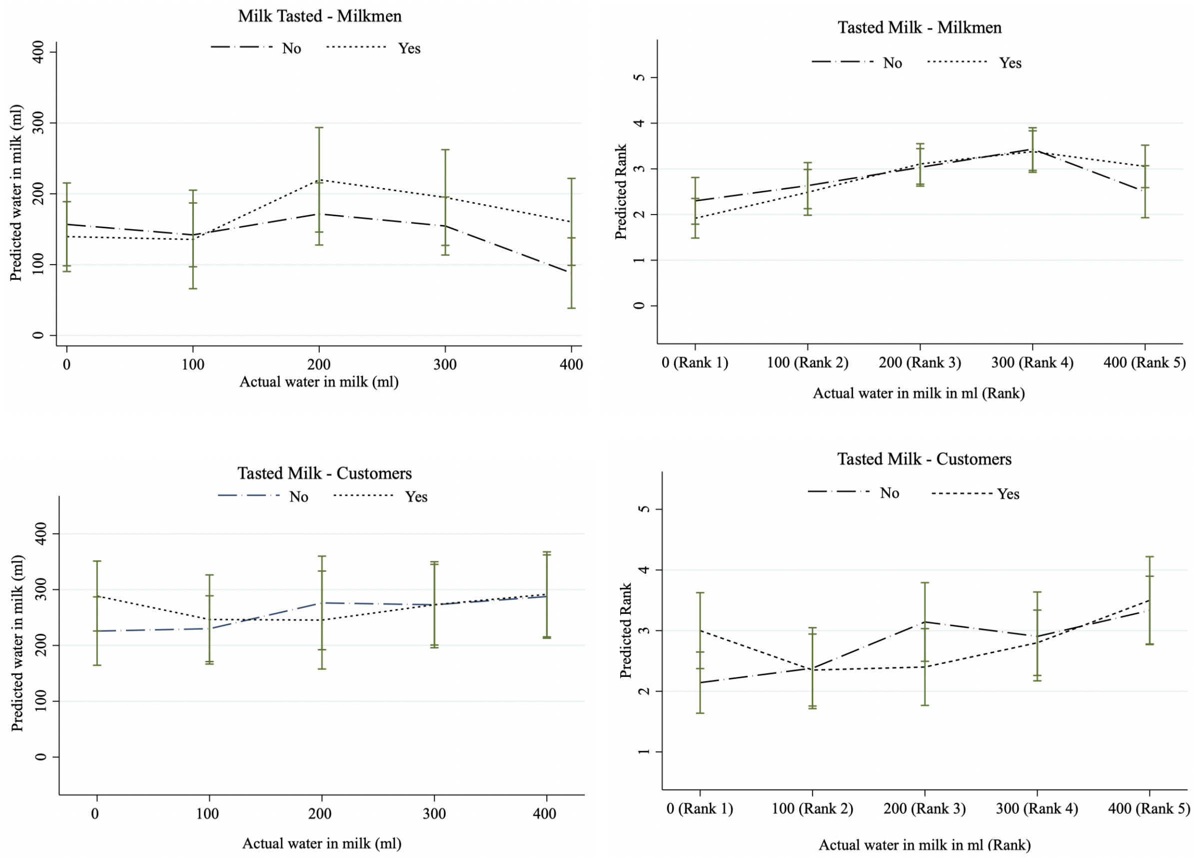


Figure A.5: Results of the milk-testing tournament by whether milk was tasted or not

Figure A.6 shows that there is no difference in predicted levels and ranks by honest and dishonest milkmen.

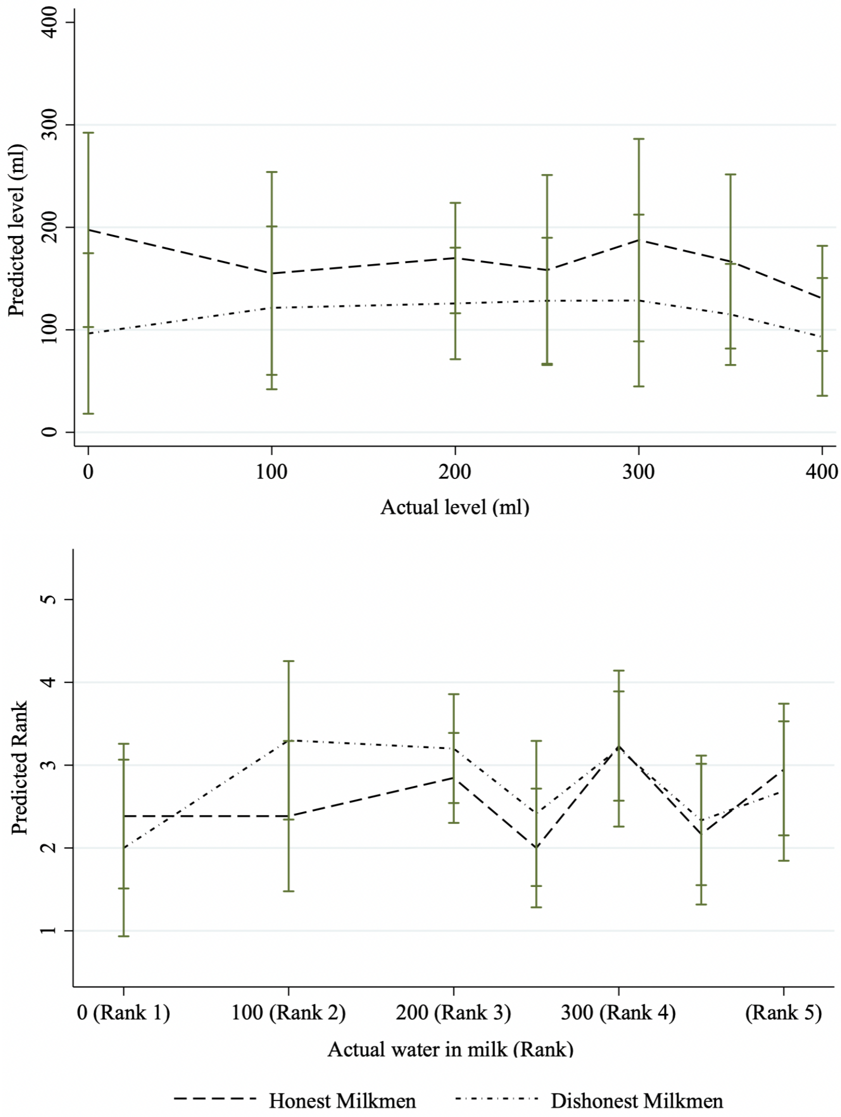


Figure A.6: Results of the milk-testing tournament by dishonesty

Appendix B

Experimental Instructions and Procedures

I. General Instructions (translated from Hindi)

Greetings and welcome to all of you. My name is XXX and his name is XXX. We are working at a university in XXX. 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 plot number 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)

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, please 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. 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 and was only readable using UV-light. We verified in the post-game interviews that milkmen did not exchange these invitation cards.

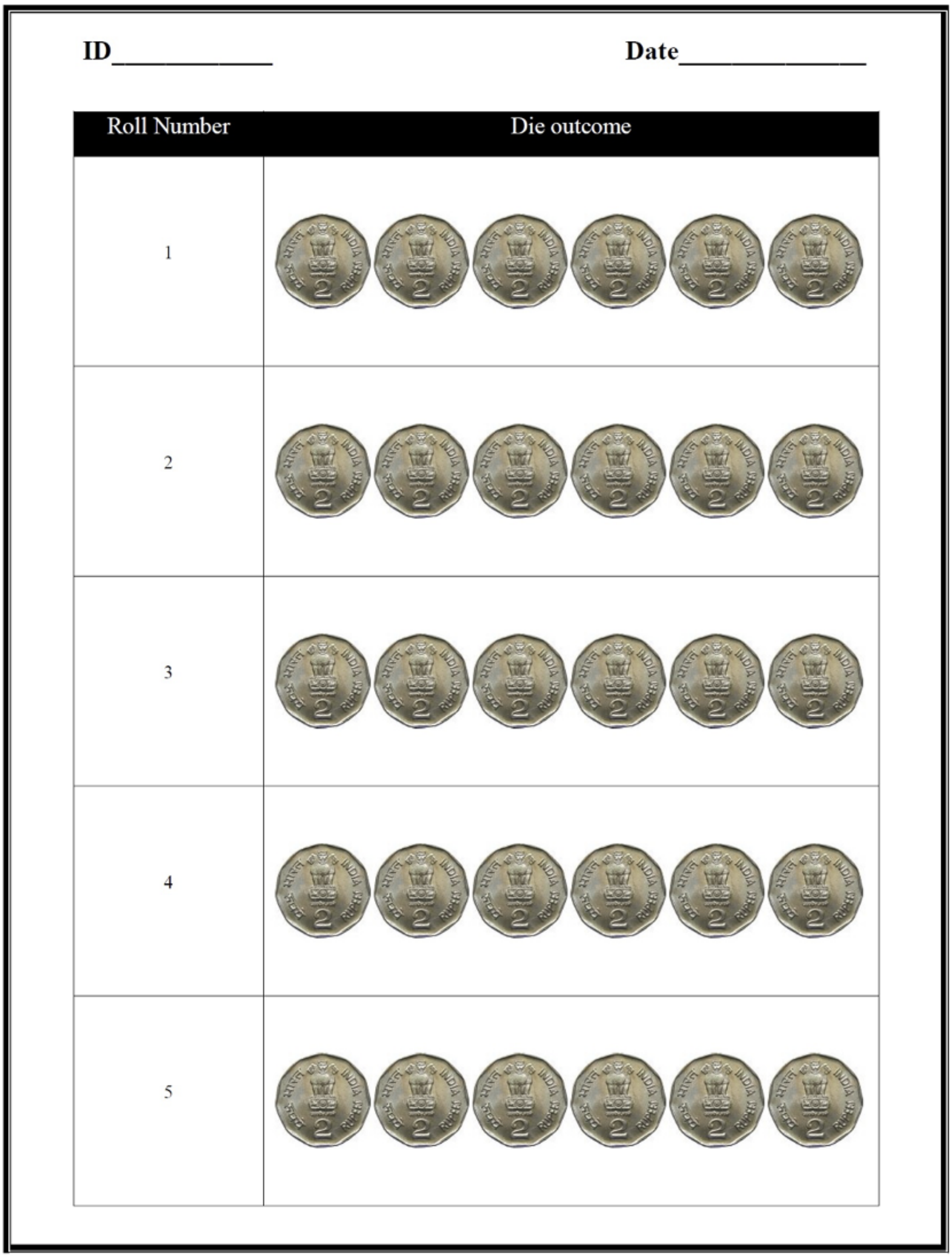
On the day of the experiment, we first carefully explained the purpose and procedure of the experiment at the group level. Each milkman then replaced his individual plot ID number with an identity card of his own choice bearing the names of European states. We then gave detailed instructions and examples 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.

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. 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 (Figure A.7). 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 A.8). 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.

Figure A.7: 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 (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.



ID____________ **Date**______________

Roll Number	Die outcome
1	
2	
3	
4	
5	

Figure A.8: Game Sheet

References

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