Insurance coverage of customers induces dishonesty of sellers in markets for credence goods

Rudolf Kerschbamer[a], Daniel Neururer[a], and Matthias Sutter[a,b,1]

[a]Department of Economics, University of Innsbruck, A-6020 Innsbruck, Austria; and [b]Department of Economics, University of Cologne, D-50923 Cologne, Germany

Edited by Jose A. Scheinkman, Columbia University, New York, NY, and approved May 17, 2016 (received for review September 9, 2015)

Honesty is a fundamental pillar for cooperation in human societies and thus for their economic welfare. However, humans do not always act in an honest way. Here, we examine how insurance coverage affects the degree of honesty in credence goods markets. Such markets are plagued by strong incentives for fraudulent behavior of sellers, resulting in estimated annual costs of billions of dollars to customers and the society as a whole. Prime examples of credence goods are all kinds of repair services, the provision of medical treatments, the sale of software programs, and the provision of taxi rides in unfamiliar cities. We examine in a natural field experiment how computer repair shops take advantage of customers' insurance for repair costs. In a control treatment, the average repair price is about EUR 70, whereas the repair bill increases by more than 80% when the service provider is informed that an insurance would reimburse the bill. Our design allows decomposing the sources of this economically impressive difference, showing that it is mainly due to the overprovision of parts and overcharging of working time. A survey among repair shops shows that the higher bills are mainly ascribed to insured customers being less likely to be concerned about minimizing costs because a third party (the insurer) pays the bill. Overall, our results strongly suggest that insurance coverage greatly increases the extent of dishonesty in important sectors of the economy with potentially huge costs to customers and whole economies.

credence goods | field experiment | insurance coverage | fraud | deception

Dishonest behavior is widespread in human societies. The frequency and degree of dishonest behavior has been shown to depend on social norms, or the lack thereof, for instance in the finance industry (1), on the costs and benefits of lying (2–4), or on the age of human subjects (5). Contrary to the effects of social or cultural norms or of personality traits, we study how insurance coverage—a key institutional arrangement in many important markets—affects the extent of dishonest behavior in markets for credence goods. The market for credence goods is huge and economically very important. Prime examples of credence goods are health care services and repair services. In the United States, for instance, health care services accounted for 17.9% of gross domestic product in 2012 (www.worldbank.org); repair services (such as for cars, office machines, and computers) are also a billion-dollar industry (6–9). Generally speaking, credence goods have the characteristic that, although customers can observe the utility they derive from the good or service ex post, they cannot judge whether the quality of the good they have received is the ex ante needed one (10–12). Moreover, customers may not even be able to observe ex post the quality they actually received. Experience goods (like wine) are related to credence goods and also have unknown characteristics, yet these characteristics are revealed after buying or consuming them, contrary to the case of credence goods (10).

The informational asymmetries between sellers and customers can generate several types of fraud (13) that bring with them large efficiency costs and have been receiving a lot of public attention. To define these systematically, consider a car owner bringing her vehicle to a garage for repair. The mechanic—as an expert seller—might have an incentive to cheat the consumer on two levels: first, the repair might be inefficient due to the mechanic replacing more parts than are actually necessary to bring the car back on the road (and charge for the additional time and material). This case is referred to as "overprovision" because the additional benefits to the consumer are smaller than the additional costs. The mechanic's repair might also be insufficient, thus leaving the consumer with a bill, but with a car that is still not running properly. This latter case is referred to as "underprovision" because any material and time spent on the repair is a pure waste. Second, the repair might be appropriate, but the mechanic might charge the consumer for more than he has actually done (e.g., by claiming to have changed a filter without having done so). This kind of problem is known as "overcharging," and it can also lead to inefficiencies in the long run if the fear of getting overcharged deters consumers from trading on credence goods markets in the future, thereby creating a type of market breakdown (14).

Insurance adds another layer of reasons for economic inefficiency in markets for credence goods. To illustrate that, consider the market for health care services and assume that the customer (a patient) is fully insured and interacts with a seller of the service (a physician). Moral hazard (15–19) implies that the patient may have incentives to demand more of the service than required, for example, by asking for more numerous or more extensive tests or treatments, because she will not bear the costs thereof. However, the behavior of the physician may also be affected by the extent of the coverage: if the physician expects the patient not to be concerned about minimizing costs, he may be more inclined to suggest or prescribe more expensive treatments (20). We are interested here in this latter source for possible inefficiencies, which we like to call "second-degree moral hazard." This term is used to highlight the fact that we are not interested in the direct effect of moral hazard which (in the case of insurance) involves an insured agent

Significance

Markets for credence goods are characterized by informational asymmetries between sellers and customers, creating strong incentives for fraudulent seller behavior. Prime examples for credence goods are repair services or medical treatments. Empirical evidence suggests that sellers' fraud imposes an excess burden of billions of dollars annually to customers. We examine in a natural field experiment how customers' insurance coverage exacerbates inefficiencies in the provision of credence goods. Specifically, we study how computer repair shops exploit customers' insurance. In a control treatment, the average repair price is about EUR 70, but increases by 80% when the customer has insurance for repair costs. We can show that the latter increase is mainly due to overprovision of parts and fraudulent overcharging of working time.

Author contributions: R.K., D.N., and M.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: matthias.sutter@wiso.uni-koeln.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518015113/-/DCSupplemental.

and an insurer, but rather in the indirect effect, which involves a third party—the supplier of a good or service who increases expenditures in ways that he or she would not in the absence of insurance coverage (21).*

Second-degree moral hazard may contribute substantially to the exploding costs in health care and repair services. However, its significance is difficult to assess with empirical data because the two phenomena (of traditional moral hazard and second-degree moral hazard) are observationally equivalent in terms of final outcomes in the sense that more extensive insurance coverage leads to higher expenditure; the mechanisms are, however, different. To investigate this in more detail, we present what has been termed a "natural field experiment" (22, 23) that allows for a clean and unambiguous measurement of the influence of second-degree moral hazard on fraud in the provision of credence goods. Natural field experiments have advantages over both empirical studies based on field data and experimental studies based on laboratory data. Compared with nonexperimental field studies, controlled variation of factors potentially important for the extent of fraud is an important advantage (24). Compared with laboratory experiments, external validity is less of an issue in a natural field experiment because participants are a representative, randomly chosen, non–self-selected subset of the treatment population of interest. Schneider (9) has pioneered natural field experiments on repair services in credence goods markets by examining how reputational concerns affect the service quality of car mechanics.

Design of the Natural Field Experiment

In our natural field experiment, we sent an undercover experimenter with manipulated test computers to 61 of the 251 registered computer repair shops all over Austria to ask for a repair. The 61 shops were randomly selected among shops along the west–east axis of Austria. The selected shops account for 24% of the total number of shops and they were reached by driving a total distance of 9,500 km (see Fig. 1 and Table S1 in the *Supporting Information* for details on randomization).

For our study, we bought five identical, completely refurbished, and perfectly working, computers (see *Methods* for the detailed specification). In each of the five computers, we destroyed one of the random access memory (RAM) modules. This manipulation prevents the computer from booting causing the following error message to appear on the screen: "ERROR 1830: Invalid memory configuration—power off and install a memory module to Slot-0 or the lower slot." Our information technology (IT) colleagues informed us that RAM modules crash from time to time, implying that the problem should be well known to experts and easy to diagnose. Moreover, a proper repair should be easily done within half an hour according to our IT department.

In the repair shop, we always used the same, fixed script, starting with the following: "When starting my computer an error message appears and I am not able to boot the computer. I have no idea what this means and I would like you to repair it, please." The second sentence was intended to create the impression of a nonexpert customer in both treatments that were implemented in a between-subject design, and with random assignment of treatments to repair shops. In the control treatment (CONTROL), before leaving the repair shop, the experimenter stated: "I will need a bill for the repair." Indicating that a bill is needed for a transaction is not uncommon in Austria, especially when dealing with relatively small businesses. It is an elegant way to communicate to the seller that the customer does not seek for a cash-in-hand job (to evade taxes). In the insurance treatment (INSURANCE), the same statement was complemented in the following way: "I will need a bill for the repair because I have insurance that covers the repair costs." Other than this minimal difference, both treatments were completely identical (we always had the same experimenter, wearing the same type of clothes, with the same computer model, the same RAM module manipulation, and the same script). These conditions allow us to identify the effects of insurance coverage on the seller's behavior. Our hypothesis is that informing the expert that the customer has insurance for the repair costs increases the amount of overprovision (more repair than needed) and overcharging (inflating bills without justification) compared with the control treatment where the customer does not indicate this.

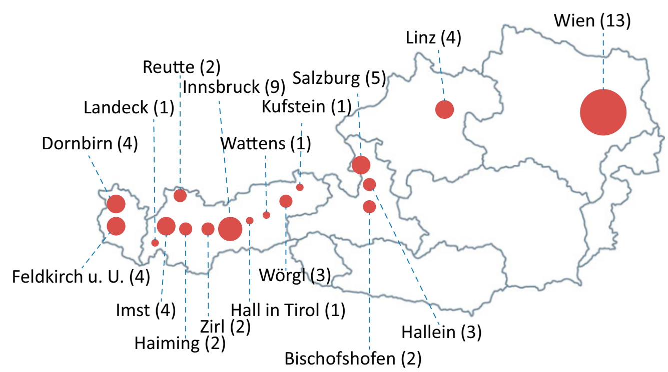

Fig. 1. Map of Austria and locations where the 61 observations were collected (number of observations per location in parentheses).

Implementation and Measurement

To be able to test our hypothesis and to interpret the results in a meaningful way, we have to keep three important issues in mind: First, it is important that the computer expert is able to diagnose the problem correctly—otherwise, measured misconduct might be due to incompetence and not intended misbehavior. The simple manipulation of the RAM module and the unambiguous error message on the screen satisfy this condition. Second, our test computers must be in perfect condition (except for our manipulation)—otherwise, an overly cautious expert might perform additional repairs not because he has material incentives for doing so but rather because he expects failure in the near future. Given that the computers were bought as completely refurbished, this condition is also met. Third, the value of the test computers should be high enough compared with repair costs; otherwise, replacement (instead of repair) was the more plausible strategy. Each computer cost EUR 684, and our IT department estimated an appropriate repair cost of EUR 60 to EUR 80, implying that repair is a plausible strategy, whereas offering a new computer is not.

The experiment was conducted between March 2013 and September 2013. The computers were handed in during regular opening hours. After a computer was picked up from a repair shop, we opened the computer and checked the repair that the shop had actually done and compared this to the positions on the bill. With these data, we are able to quantify overprovision, overcharging in the spare-parts dimension, and overcharging in the working-time dimension. Any repair not related to the replacement of the broken RAM module is considered as overprovision because the test computers were in a perfect condition apart from the manipulated RAM module. Overcharging in the spare-parts dimension refers to cases where repairs were listed on the bill, but not actually done. Overcharging in the working-time dimension is measured by

*Second-degree moral hazard has been studied in highly regulated markets, like the market for taxi rides (21). However, in highly regulated markets the scope of fraudulent behavior is limited, and the different sources for fraudulent behavior (overtreatment and overcharging) have different consequences for customers in the market for taxi rides (with overtreatment putting a burden on the time spent in the taxi, while overcharging has only financial consequences). In our natural field experiment, both dimensions (overtreatment and overcharging) are ultimately part of the bill, and we can disentangle which dimension contributes which fraction.

comparing the indicated working time between the control treatment and the insurance treatment.

Results

In total, 58 of the 61 repair shops were successful with the repair. The remaining three shops claimed that the computer was not repairable or that a repair made no sense because it would be more expensive than buying a new computer. One of these three shops was in CONTROL, and this shop did the diagnosis for free. The other two shops were in INSURANCE, and they charged EUR 57 and EUR 80, respectively, for the diagnosis. These three cases of unsuccessful repair are excluded from the following.

The mean duration from handing in the computer to the call that the laptop is ready for pickup was 2.07 d. The 95% confidence interval (CI) was ±0.62 d, and the median was 1 d. The average repair price in CONTROL was EUR 70.17 (CI: ±12.54; median: 70.00) and in INSURANCE it was EUR 128.68 (CI: ±29.10; median: 95.4). The difference is economically impressive and statistically highly significant (Mann–Whitney test: $P = 0.0015$; $n = 58$). This nonparametric result is corroborated in two regressions—shown in Tables S2 and S3—where we control for the city, for the presence of office premises, for the fact whether a shop only repairs computers or also sells them, and whether the shop is a one-man business. The estimated effect of the INSURANCE-treatment is about EUR 60. Fig. 2 shows the relative cumulative frequencies of total repair prices. It shows clearly the large difference between INSURANCE and CONTROL.

We now turn to the sources of the large treatment difference. We count five observations with additional repairs not related to our manipulation and all of them took place in INSURANCE (with a total of 27 observations). Because we ensured that our test computers were in perfect condition (except for our RAM manipulation), we classify these additional, but unnecessary, repairs as overprovision. The Fisher's exact test reveals that overprovision is significantly more frequent in INSURANCE than in CONTROL (17.24% vs. 0%; $P = 0.018$; $n = 58$). In all five cases, the broken RAM module was replaced (otherwise, the laptop would not work at all) and the other repairs were conducted in addition, yielding an average repair cost of EUR 200.58 (CI: ± 86.21; median: 222.00). The costs in these five cases are significantly higher than the average repair price in the other 22 observations in INSURANCE (which was EUR 112.34; CI: ± 28.97; median: 87.17; Mann–Whitney test: $P = 0.0227$; $n = 27$). Hence, overprovision explains one part of the price difference between CONTROL and INSURANCE. The other part must be related to overcharging, which can happen in the spare-parts dimension (billing for spare parts that have not been provided) or the working-time dimension

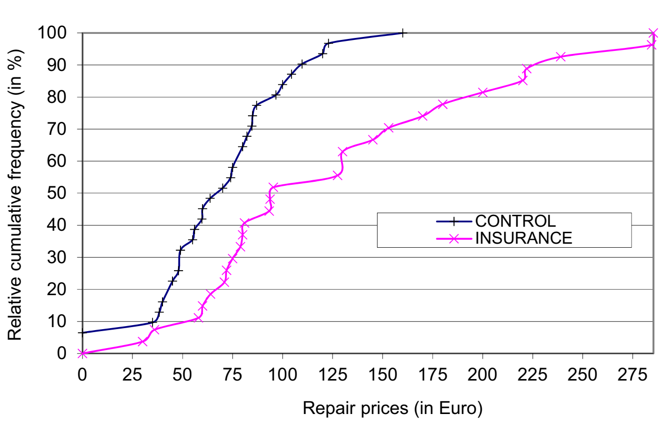

Fig. 2. Relative cumulative frequency of repair prices (in euros), conditional on treatment ($n = 58$).

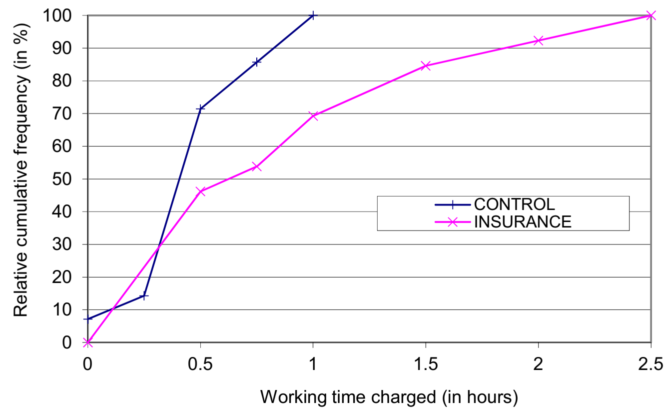

Fig. 3. Relative cumulative frequency of working time (in hours), conditional on treatment ($n = 27$, excluding observations with overtreatment).

(billing for more repair time than is obviously necessary in the CONTROL treatment).

We find no treatment difference in overcharging in the spare-parts dimension. There are four cases of billing for a spare part that was not actually provided (which we verified by opening the computer and checking for replacements): two of these cases were in CONTROL, and the other two in INSURANCE.

Overcharging in the working-time dimension is more difficult to detect because the undercover experimenter was not present during diagnosis and repair. Here, we proceed as follows: Of the 58 repair shops with successful repair, 29 indicated the working time and hourly rate on the bill. (Austrian law does not prescribe that working time and hourly rate are listed individually on the bill. Despite this, one-half of the shops displayed the respective information, with no treatment difference in the likelihood of adding this information on the bill.) From the 29 shops who indicated the working time and hourly rate on the bill, two were excluded from the following analysis because these two cases were among the five cases of overtreatment (and because repairing additional parts arguably requires more time). Among the remaining 27 observations, we note an average working time of 0.55 h in CONTROL (CI: ± 0.14; median: 0.5), but of 1.02 h in INSURANCE (CI: ± 0.40; median: 0.75; Mann–Whitney test: $P = 0.0463$; $n = 27$)[†] (25). Fig. 3 shows the relative cumulative frequency of working time, again indicating a strong difference between CONTROL and INSURANCE.

Given average hourly rates of EUR 87.47 (CI: ± 7.79; median: 87.6)—with no significant treatment difference—the average difference of 0.47 h (28 min) in working time between INSURANCE and CONTROL increases the total bill by about EUR 41.11, which accounts for roughly 70% of the overall repair price difference of EUR 58. The remaining 30% is due to overtreatment by performing unnecessary repairs.

To find out more about the motives for the discriminatory behavior of sellers in our natural field experiment, we conducted a survey in November 2015 in 15 repair shops (about one-half of them served also as subjects in our main natural field experiment more than 2 y earlier) where we asked the persons at the front desk about why insurance coverage might lead to higher prices (see *Supporting Information* for the questionnaire and for further results). We told

[†]To check whether multiple testing has a notable influence on our results, we use the recently developed method of List et al. (25). Simultaneously testing for repair prices, overprovision, and working time (our main results), we get the following results. Average repair price CONTROL vs. INSURANCE: $P = 0.004$; average working time CONTROL vs. INSURANCE: $P = 0.058$; overprovision CONTROL vs. INSURANCE: $P = 0.097$. If we subsume overprovision, overcharging in working time and in spare parts under the category "misbehavior" and test simultaneously for the significant differences in repair prices and in misbehavior, we get $P = 0.004$ for repair prices and $P = 0.016$ for misbehavior.

Table 1. Mean answers and SD for each of the five questions

Question	Q1 (prudence)	Q2 (second-degree moral hazard)	Q3 (animus-based discrimination)	Q4 (third-degree price discrimination)	Q5 (misperceived third-degree price discrimination)
Mean	3.4	1.73	2.47	2.73	3.13
Standard deviation (SD)	0.828	0.704	0.743	0.798	0.833

Answers range from 1 (very likely) to 4 (very unlikely).

the interviewees first that scientific studies had shown that insured customers face higher bills for repair services, and then we asked them to indicate on a scale from 1 (very likely) to 4 (very unlikely) to assess the possible reasons for this finding. Our questions were intended to discriminate between the following explanations for the observed treatment differences: (Q1) prudence: experts invest more time and effort when the computer is insured because they feel solidarity with customers and intend to help them; (Q2) second-degree moral-hazard: experts charge higher prices when the computer is insured because insured customers are perceived as being less concerned about cost minimization because a third party (the insurer) pays the bill; (Q3) animus-based discrimination: experts charge higher prices when the computer is insured because insurance companies are seen as leeches that cause harm to society and are therefore treated worse than normal customers (26); (Q4) classical third-degree price discrimination: experts charge higher prices when the computer is insured because insurance companies are perceived as rich and because experts tend to charge higher prices from richer customers (because they expect that richer customers have a higher willingness to pay); and (Q5) misperceived third-degree price discrimination: experts charge higher prices when the computer is insured because insured customers are perceived as having a higher willingness to pay not because they do not pay the bill themselves (this is the explanation in Q2) but rather because by buying insurance they have revealed that the computer is important for them.

Among the five alternatives, responders answered most affirmative to the following potential explanation (Q2): "Due to the insurance coverage, the customer has no incentive to pay attention to the costs. Expert sellers are aware of this and therefore charge higher prices." The average response here is 1.73 (CI: ± 0.39; median: 2), which is significantly lower (and thus considered as more likely) than for all other potential explanations. See Table 1 and Table S4 for details; the Wilcoxon signed-ranks test yields $P < 0.01$ in all cases (*Supporting Information*). Hence, the results from the survey suggest that insurance coverage is not primarily perceived as signal of higher willingness to pay, but rather as an indicator that the customer is less likely to be concerned about cost minimization because a third party (the insurer) covers the cost.

Discussion

Insurance services provide the potential for moral hazard problems between the insured agent and the insurance company, as economists have documented for 40 y. We have shown that insurance can have an additional effect on the behavior of an insured agent's transaction partner, when that partner knows that the costs of his misbehavior are covered by the insurer. This phenomenon—which we like to call second-degree moral hazard—has widespread implications especially for markets plagued by informational asymmetries between expert sellers and customers. This observation adds to our understanding of motivations for honest vs. dishonest behavior in economic transactions.

Our findings for the computer repair market indicate that second-degree moral hazard can inflate bills through fraudulent overprovision and overcharging to a considerable extent, in our case up to about 80%, on average. The damage done to insurers is sizable and we would expect this to lead to higher insurance premiums for customers. Furthermore, the economy as a whole also pays a price through the unnecessary overprovision of goods, which is a waste of scarce resources. Not coincidentally, many insurers have chosen to develop long-term relationships with specific repair businesses—physicians in the case of health insurance—because long-term relationships (and the threat of their termination) might serve to curtail second-degree moral hazard.

Methods

The experiment was approved by the Internal Review Board of the University of Innsbruck. There was no debriefing or any other kind of information to subjects (before or after the experiment) that revealed that the shops that we visited were part of an experiment.

Selection of Shops. First, we compiled a list of all repair shops located in the geographic area of interest and assigned to each shop on the list a specific number. Then, we used a random number generator (implemented in Wolfram's Mathematica) to rank the shops, and picked the first 61 shops. Then we randomly allocated to each of these shops one of our treatments (INSURANCE or CONTROL). A list of shops (251 in total) is available in the Austrian telephone directory (HEROLD, www.herold.at/gelbe-seiten/computer-reparatur-u-service/).

Specification of Notebooks. We bought five identical notebooks to be able to speed up the data collection process by bringing each of them to a different shop. The notebooks were completely refurbished and the cost was EUR 684 for each. The notebooks had the following configuration: NTR Lenovo TP T500 2089-A35 refurbished/Intel Core 2 Duo P8600 Processor 2 × 2, 40 GHz/4,096-MB RAM/160-GB hard disk drive/15′ 4″ 1,680 × 1,050 pixel (WSXGA + TFT) flat display/Intel Graphics Media Accelerator ×4,500HD/DVD burner/WLAN a/b/g/n/Bluetooth/UMTS integrated/Windows 7 Professional (64 bit)/12-mo warranty/new battery.

Manipulation of Notebooks. After the purchase of the notebooks, we personalized them by storing folders and personalized files on each of them. Then we opened each computer and broke one of the two RAM modules. This was done by mechanically destroying a small chip on each RAM module with the consequence that the whole RAM module did not work anymore. This manipulation was not visible because it was very small and because we hid it underneath the company sticker of the RAM module. Then we closed the notebook and brought it to a repair shop.

ACKNOWLEDGMENTS. We thank two anonymous reviewers, Andreas Lange, and Andreas Nicklisch, and seminar audiences in Cologne, Hamburg, Innsbruck, London (Royal Holloway), Oxford, Rotterdam, Vienna (University of Economics and Business Administration), and Warwick (Business School) for helpful comments. Katherine Maye-Saidi was of great help in editing the text. Financial support from the Vice Rectorship of Research of the University of Innsbruck through Grant 2013/1/VWL21 and from the Austrian Science Fund through Grant P 26901-G11 is gratefully acknowledged.

1. Cohn A, Fehr E, Maréchal MA (2014) Business culture and dishonesty in the banking industry. *Nature* 516(7529):86–89.
2. Gneezy U (2005) Deception: The role of consequences. *Am Econ Rev* 95(1):384–394.
3. Sutter M (2009) Deception through telling the truth?! Experimental evidence from individuals and teams. *Econ J* 119(534):47–60.
4. Erat S, Gneezy U (2012) White lies. *Manage Sci* 58(4):723–733.
5. Glätzle-Rützler D, Lergetporer P (2015) Lying and age: An experimental study. *J Econ Psychol* 46:12–25.
6. Wolinsky A (1993) Competition in a market for informed experts' services. *Rand J Econ* 24(3):380–398.
7. Wolinsky A (1995) Competition in a market for credence goods. *J Inst Theor Econ* 151(1): 117–131.
8. Gruber J, Kim J, Mayzlin D (1999) Physician fees and procedure intensity: The case of cesarean delivery. *J Health Econ* 18(4):473–490.
9. Schneider HS (2012) Agency problems and reputation in expert services: Evidence from auto repair. *J Ind Econ* 60(3):406–433.
10. Darby M, Karni E (1973) Free competition and the optimal amount of fraud. *J Law Econ* 16(1):67–88.
11. Dulleck U, Kerschbamer R (2006) On doctors, mechanics and computer specialists: The economics of credence goods. *J Econ Lit* 44(1):5–42.

12. Dulleck U, Kerschbamer R, Sutter M (2011) The economics of credence goods: On the role of liability, verifiability, reputation and competition. *Am Econ Rev* 101(2): 526–555.
13. Balafoutas L, Beck A, Kerschbamer R, Sutter M (2013) What drives taxi drivers? A field experiment on fraud in a market for credence goods. *Rev Econ Stud* 80(3):876–891.
14. Akerlof G (1970) The market for lemons: Quality uncertainty and the market mechanism. *Q J Econ* 84(3):488–500.
15. Holmstrom B (1979) Moral hazard and observability. *Bell J Econ* 10(1):74–91.
16. Hubbard T (1998) An empirical examination of moral hazard in the vehicle inspection market. *Rand J Econ* 29(2):406–426.
17. Cason T, Gangadharan L, Maitra P (2012) Moral hazard and peer monitoring in a laboratory microfinance experiment. *J Econ Behav Organ* 82(1):192–209.
18. Huck S, Kübler D, Weibull J (2012) Social norms and economic incentives in firms. *J Econ Behav Organ* 83(2):173–185.
19. Füllbrunn S, Neugebauer T (2013) Limited liability, moral hazard, and risk taking: A safety net game experiment. *Econ Inq* 51(2):1389–1403.
20. Huck S, Lünser G, Spitzer F, Tyran J-R (2014) *Medical Insurance and Free Choice of Physician Shape Patient Overtreatment. A Laboratory Experiment* (Department of Economics, University of Copenhagen, Copenhagen), Discussion Paper No. 14-19. Available at https://www.econstor.eu/bitstream/10419/103210/1/798554401.pdf. Accessed June 3, 2016.
21. Balafoutas L, Kerschbamer R, Sutter M (2015) Second degree moral hazard in a credence goods market. *Econ J*, in press.
22. Harrison G, List JA (2004) Field experiments. *J Econ Lit* 42(4):1009–1055.
23. Levitt SD, List JA (2009) Field experiments in economics: The past, the present, and the future. *Eur Econ Rev* 53(1):1–18.
24. Zitzewitz E (2012) Forensic economics. *J Econ Lit* 50(3):731–769.
25. List JA, Shaikh AM, Xu Y (2016) Multiple hypothesis testing in experimental economics. NBER working paper 21875. Available at home.uchicago.edu/amshaikh/webfiles/experimental.pdf. Accessed June 3, 2016.
26. Gneezy U, List JA, Price MK (2012) Toward an understanding of why people discriminate: evidence from a series of natural field experiments. NBER working paper 17855. Available at www.nber.org/papers/w17855.pdf. Accessed June 3, 2016.

Supporting Information

Kerschbamer et al. 10.1073/pnas.1518015113

Additional Analysis

For each observation, we have the following data: hand-in date; date of the day when the repair was done; pickup date; address of the repair shop; and whether the shop was a one-man or a multiagent shop, whether it had office premises or not, and whether it conducted just repairs or also sold computers. Most importantly, we have the total repair price for all observations, and for about one-half of them (29) we have in addition the exact working time and the price for it, as well as the used spare parts and their price.

There is no significant relationship between the treatment (INSURANCE and CONTROL) and the fact that there is detailed information (working time and price per hour) on the repair bill (χ^2 test: $P = 0.992$; $n = 56$). [Of the 58 shops that successfully repaired the computer, two shops (both in CONTROL) did not charge anything, and they are excluded in the analysis in this paragraph.] The average total repair price is also not significantly different when there is no detailed information on the repair bill: EUR 107 (median: 80; SD: 68.14; 95% CI: 81.44, 133.28) vs. EUR 94 (median: 81; SD: 52.98; 95% CI: 72.98, 114.90). Here, the Mann–Whitney test yields $P = 0.7804$ ($n = 56$). The average duration until the shop informed us that the pickup is ready is significantly higher when there is no detailed information on the repair bill: 2.7 d (median: 2; SD: 2.45; 95% CI: 1.73, 3.63) vs. 1.6 d (median: 1; SD: 2.12; 95% CI: 0.72, 2.40). Here, the Mann–Whitney test yields $P = 0.0292$ ($n = 55$). (We had to drop one observation because the shop forgot to call us when the notebook was ready for pickup.) When we have the information on hourly rates for repair, there is no treatment difference between INSURANCE and CONTROL: EUR 85 (median: 82; SD: 19.60; 95% CI: 73.69, 95.40) vs. EUR 91 (median: 91.2; SD: 21.65; 95% CI: 78.09, 103.09); Mann–Whitney test: $P = 0.4845$ ($n = 29$). However, there is a treatment difference with respect to the average duration until pickup was ready, with an average of 1.6 d (median: 1; SD: 2.04; 95% CI: 0.77, 2.36) in CONTROL and 2.7 d (median: 2; SD: 2.52; 95% CI: 1.70, 3.70) in INSURANCE (Mann–Whitney test: $P = 0.0477$; $n = 58$).

Balancing Tests to Check Whether the Randomization Was Successful. See Table S1.

Regression Analysis: What Determines Repair Prices and Working Time Charged? See Table S2.

Regression Analysis: What Determines Repair Prices and Working Time Charged? (Including Location Dummies). See Table S3.

Survey Among Repair Shops

The following survey was conducted in 15 shops in November 2015. All shops are located in Innsbruck and the surrounding area. We have visited all repair shops with office premises in this area. We have chosen to conduct the survey in Innsbruck and the surrounding area because it is plausible that the University of Innsbruck conducts a survey in this area. Furthermore, about one-half of the survey participants served also as subjects in the main experiment more than 2 y earlier.

Translation of Survey (from German). My name is Daniel Neururer, and I am a research assistant at the University of Innsbruck. Within the framework of my research about markets for repair services, I stumbled across a question. By giving me your opinion on an explanation for this question, you would help me greatly.

In some scientific studies, there are data suggesting that customers who have insurance that covers repairs pay more on average for the same repair than customers who cover the costs of repairs themselves.

Why do you think that some experts charge customers with insurance coverage higher prices than customers without insurance coverage?

Please evaluate each of these potential explanations according to their degree of probability:

1 = very likely; 2 = likely; 3= unlikely; 4 = very unlikely.

- Q1: Due to the insurance coverage, the repair is done more thoroughly and therefore takes longer. The costs are increased because of this.
  Evaluation:__________
- Q2: Due to the insurance coverage, the customer has no incentive to pay attention to the costs. Expert sellers are aware of this and therefore charge higher prices.
  Evaluation:__________
- Q3: Insurance companies have a very bad reputation. For this reason, experts charge comparatively more for repairs covered by insurance because in doing so they only cause harm to the insurance company but not to the customer.
  Evaluation:__________
- Q4: Insurance companies generally have the reputation that they are swimming in money. For this reason, experts charge customers with insurance higher prices as they generally tend to charge wealthy customers more for repairs.
  Evaluation:__________
- Q5: Optional insurance coverage signals that the repair is very important for the customer and that he is willing to pay more. The experts exploit this by charging comparatively more for repairs.
  Evaluation:__________

Results of the Survey. The results of the survey are displayed in Table S4. In the table, Qx refers to question x in the above translated questionnaire.

Table S1. Table of preexperiment observable characteristics of the repair shops

Observable characteristics	CONTROL	INSURANCE	Test statistic
Proportion of shops that sell computers	0.71	0.67	Fisher's exact test: $P = 0.781$; $n = 58$
Proportion of shops with office premises	0.81	0.85	Fisher's exact test: $P = 0.737$; $n = 58$
Proportion of shops located in a city	0.68	0.67	Fisher's exact test: $P = 1.000$; $n = 58$
Proportion of one-man shops	0.35	0.44	Fisher's exact test: $P = 0.593$; $n = 58$

Table S2. OLS regression 1

Dependent variable (OLS regressions)	Repair price, EUR	Working time, h
Variable		
INSURANCE-treatment (1 = yes)	59.82***	0.46**
	(15.08)	(0.19)
City dummy (1 = city)	−6.23	−0.17
	(18.26)	(0.20)
Office premises (1 = yes)	−27.70	−0.21
	(21.69)	(0.24)
Selling computers (1 = yes)	−10.67	0.09
	(18.61)	(0.27)
One-man shop (1 = yes)	−6.45	0.29
	(19.26)	(0.24)
Constant	106.59***	0.60*
	(30.87)	(0.35)
No. observations	58	27

OLS (ordinary least squares) regression with repair price (in euros) and working time (in hours) as independent variables, including as explanatory variables a treatment dummy for INSURANCE, a dummy for the shop being in a city, whether the shop has office premises, whether it also sells computers (and not only repairs computers), and whether the shop is a one-man business. ***, **, and * denote significance at the 1%, 5%, 10% levels. SEs are in parentheses.

Table S3. OLS regression 2

Dependent variable (OLS regressions)	Repair price, EUR	Working time, h
Variable		
INSURANCE-treatment (1 = yes)	64.23***	0.67***
	(16.33)	(0.23)
Office premises (1 = yes)	−13.97	0.18
	(23.91)	(0.34)
Selling computers (1 = yes)	−8.64	−0.09
	(18.36)	(0.31)
One-man shop (1 = yes)	−3.30	0.42
	(17.97)	(0.26)
Located in Tirol (1 = yes)	14.02	0.58
	(20.46)	(0.32)
Located in Vorarlberg (1 = yes)	−23.28	0.27
	(26.93)	(0.37)
Located in Salzburg (1 = yes)	30.79	0.79
	(25.16)	(0.52)
Located in Oberösterreich (1 = yes)	17.07	0.20
	(34.37)	(0.56)
Constant	78.07**	−0.21
	(32.36)	(0.50)
No. observations	58	27

OLS (ordinary least squares) regression with repair price (in euros) and working time (in hours) as independent variables, including as explanatory variables a treatment dummy for INSURANCE, a dummy for whether the shop has office premises, whether it also sells computers (and not only repairs computers), dummies for the different provinces (Vienna is the reference group), and whether the shop is a one-man business. ***, **, and * denote significance at the 1%, 5%, 10% levels; SEs are in parentheses.

Table S4. Mean responses to survey among repair shops (run in November 2015)

Questions	Q1	Q2	Q3	Q4	Q5
Q1 (prudence)	3.4 [0.828] {15}				
Q2 (second-degree moral hazard)	$P = 0.001$***	1.73 [0.704] {15}			
Q3 (animus-based discrimination)	$P = 0.015$**	$P = 0.003$***	2.47 [0.743] {15}		
Q4 (third-degree price discrimination)	$P = 0.071$*	$P = 0.001$***	$P = 0.103$	2.73 [0.798] {15}	
Q5 (misperceived third-degree price discrimination)	$P = 0.103$	$P = 0.001$***	$P = 0.026$**	$P = 0.211$	3.13 [0.833] {15}

Diagonal: mean. Standard deviation is in brackets, and number of observations is in braces. Below the diagonal: *P* value of the Wilcoxon signed-rank test. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.