UNIVERSITÉ DE
NEUCHÂTEL

FACULTÉ DES SCIENCES
ÉCONOMIQUES

Theory and evidence on the credence component of energy-efficient technologies

PhD thesis submitted to the Faculty of Economics and Business
Institute of Economic Research
University of Neuchâtel

For the PhD degree in Economics
by

Evert REINS

Approved by the dissertation committee:

Prof Bruno Lanz, University of Neuchâtel, thesis director
Prof. Mehdi Farsi, University of Neuchâtel
Prof. Massimo Filippini, ETH Zürich
Prof. Timo Goeschl, University of Heidelberg, Germany
Prof. Milad Zarin, University of Neuchâtel

Defended on September 22, 2021

Faculté des sciences économiques
Avenue du 1er-Mars 26
CH-2000 Neuchâtel
www.unine.ch/seco

LETTRES ET SCIENCES HUMAINES
SCIENCES
DROIT
SCIENCES ECONOMIQUES

Contact:
Evert Reins
evert.reins@unine.ch
University of Neuchâtel, Institute of Economic Research,
Rue A.-L.Breguet 2, 2000 Neuchâtel, Switzerland

UNIVERSITÉ DE
NEUCHÂTEL

IMPRIMATUR POUR LA THÈSE

Theory and evidence on the credence component of
energy-efficient technologies

Evert REINS

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

La Faculté des sciences économiques,
sur le rapport des membres du jury

Prof. Bruno Lanz (directeur de thèse, Université de Neuchâtel)
Prof. Milad Zarin (président du jury, Université de Neuchâtel)
Prof. Massimo Filippini (ETH Zürich)
Prof. Timo Goeschl, Université de Heidelberg

Autorise l'impression de la présente thèse.

Neuchâtel, le 23 novembre 2021

Le doyen
Valéry Bezençon

To my mother, Sabine.

Contents

List of Figures

Chapter II **33**

Chapter III **63**

Appendix **97**

List of Tables

Résumé

La composante de confiance des technologies à haut rendement énergétique constitue un obstacle important à leur adoption. Dans le contexte de nombreux investissements dans les technologies à haut rendement énergétique, la demande du marché dispose d'informations limitées sur les technologies ou les services optimaux et doit donc faire confiance à la cote de l'offre du marché. Dans ce cas, l'offre peut exploiter son avantage informationnel, ce qui conduit à des inefficacités caractéristiques des biens de confiance, telles que la fourniture de services insuffisants/extensifs ou la surfacturation de ces services. Dans le premier chapitre de cette thèse, je développe un modèle théorique simple mettant en évidence les inefficacités résultant de la structure d'incitation associée avec les biens de confiance. En discutant des liens entre la littérature empirique sur les biens de confiance et celle sur le marché de l'efficacité énergétique, j'identifie des implications pour la conception de politiques promouvant l'adoption des technologies à haut rendement énergétique. Dans le deuxième chapitre, j'utilise des données sur les systèmes solaires photovoltaïques (PV) subventionnés pour étudier l'aléa moral de second degré (c'est-à-dire les installateurs augmentent les inefficacités lorsque les consommateurs reçoivent des subventions). En utilisant une spécification de variable instrumentale, je quantifie l'impact des niveaux de subvention sur la production d'électricité attendue autodéclarée («design factor») qui influence le total des subventions reçues et les prix de transaction des systèmes PV. Les résultats montrent une association significative et positive entre les niveaux de subvention plus élevés et le «design factor» ainsi que les prix de transaction. Dans le troisième chapitre, j'analyse comment la réciprocité affecte les inefficacités associée avec les biens de confiance. En utilisant un cadre expérimental standard pour les expériences en laboratoire sur les biens de confiance, j'étudie comment l'échange de cadeaux modifie le comportement des consommateurs et des vendeurs experts. Les résultats suggèrent que des cadeaux modestes et inconditionnels peuvent améliorer les résultats du marché, en particulier lorsque les consommateurs ont besoin d'un service coûteux.

Mots-clés: Efficacité énergétique; Biens de confiance; Information asymétrique; Aléa moral de second degré ; Subventions; Réciprocité; Échange de cadeaux.

JEL classification: C91: D82; H23; Q42; Q55; Q58

Abstract

The credence component of energy-efficient technologies is an important barrier to their adoption. In the context of many investments in energy-efficient technologies, the demand-side of the market has limited information about which technology or services are optimal and therefore has to trust the supply-side. In this setting, the supply-side may exploit their information advantage, leading to supply-side inefficiencies characteristic for credence goods, such as providing insufficient/extensive services as well as overcharging for these services. In the first chapter of this thesis, I develop a simple theoretical framework highlighting inefficiencies resulting from the associated incentive structure. Discussing linkages between the empirical literature on credence goods and that on the market for energy efficiency, I identify implications for the design of policies promoting the adoption of energy-efficient technologies. In the second chapter, I use data on subsidized solar photovoltaic (PV) systems to study second-degree moral hazard (i.e. the impulse of installers to increase supply-side inefficiencies when consumers receive subsidies). Using an instrumental variable strategy, I quantify the impact of subsidy levels on the self-reported expected electricity output (design factor) influencing the total subsidies received and transaction prices of PV systems. The results show a significant and positive association between larger subsidy levels and the design factor as well as transaction prices. In the third chapter, I analyze how reciprocity affects supply-side inefficiencies. Employing a standard experimental framework for lab-experiments on credence goods, I study how gift exchange changes the behavior of consumers and expert sellers. Results suggest that small and unconditional gifts may improve market outcomes, in particular when consumers need an expensive service.

Keywords: Energy efficiency; Credence goods; Asymmetric information; Second-degree moral hazard; Subsidies; Reciprocity; Gift exchange.

JEL classification: C91: D82; H23; Q42; Q55; Q58

Acknowledgments

First and foremost, I want to thank my supervisor Prof. Bruno Lanz for his great support during the course of my PhD. He gave me the freedom to explore on my own, and at the same time provided the guidance and feedback I needed. I appreciate all his contributions of time, ideas and funding which made this writing experience productive and stimulating.[1] His passion and enthusiasm about research have made the last four years a great experience where I learned a lot. I am also grateful to my PhD Committee with Prof. Mehdi Farsi, Prof. Massimo Filippini, Prof. Timo Goeschl, and Prof. Milad Zarin, for helping me improve the thesis.

I would further like to thank Winand Emons, François Maréchal, Philippe Thalmann, Jordi Brandts, Marco Casari, Winand Emons, Holger Herz, Nicola Mauri, John Moore, Klaus Schmidt, Octavian Strimbu, Jörgen Weibull, Christian Zehnder and numerous anonymous referees for useful comments and discussions. I want to give a special thank you to Serhiy Kandul, who has been a great source of inspiration in setting up the experimental work of this thesis. I am also indebted to my friend and colleague Frank Schlütter for his intellectual support and for showing me the quickest checkmate in online chess that I have so far witnessed.

In addition, I would like to give shoutouts to my colleagues from the University of Neuchâtel and the Study Center Gerzensee (cohort of 2018/19). Without you, my time in Switzerland would not have been half as fun.

Finally, I am incredibly grateful to my girlfriend Debbie, my family and friends for providing all the patience and distractions that were necessary to finish this thesis, and for always helping me remember what is most important. I owe a special thanks to my sister Leonie, my brother in law Dylan and my dad Roelf-Evert for their precious advice

---

[1] I also gratefully acknowledge financial support from the Swiss National Science Foundation under grant PYAPP1_173650. This research is further part of the activities of SCCER-CREST (Swiss Competence Center for Energy Research), which is financially supported by Innosuisse.

and for always believing in me.

Introduction

The adoption of energy-efficient technologies plays a key role in achieving energy policy goals of many countries (see for example Alberini et al. (2016) for a discussion of fuel economy of vehicles in Switzerland and Allcott and Greenstone (2012) for a general discussion). One important barrier to the adoption of energy-efficient technologies is their credence component, implying that informational asymmetries between the demand- and the supply-side of the market may lead to supply-side inefficiencies including overcharging for services, recommending extensive services, and reducing labor input (see for example Giraudet et al., 2018; Lanz and Reins, 2021). In particular, the demand-side (or the consumer) of the market has typically limited knowledge about which technologies or services are needed to make optimal investments (Giraudet, 2020). Furthermore, it is difficult to compare the energy-efficiency of technologies after the investment, because a counterfactual measure for alternative technologies is missing. Therefore, the demand-side needs to trust the supply-side (or the expert seller) to recommend and install suited technology for a fair price (Emons, 1997; Dulleck and Kerschbamer, 2006).

The three chapters of this thesis are centered around market inefficiencies caused by the credence component of energy-efficient technologies. In the first chapter, I establish a simple framework linking a model of investments in energy-efficient technologies (Allcott and Greenstone, 2012) to the general framework of credence goods (Dulleck and Kerschbamer, 2006). This framework enables me to conduct a systematic literature survey where I identify drivers of supply-side inefficiencies for energy-efficient technologies and how they may prevent consumers from interacting causing inefficiently low investments. In the second chapter, I document empirical evidence that third-party reimbursements such as subsidies increase inefficiencies related to solar photovoltaic systems. In the third chapter, I study how gift exchange and reciprocal expert sellers affect inefficiencies and provide evidence that gifting expert sellers can improve efficiency of credence goods markets.

The ecological system of energy-efficient technologies is vast and the credence component of energy efficiency may affect particular technologies to a higher extent. Table 1 categorizes energy-efficient technologies along two domains. The first domain is the intended use of technologies. Solar photovoltaic systems, cars and fridges all have in common that they transform one source of energy into another (see column 1 of Table 1). In this context, efficient energy-transforming technologies should have a high ratio of energy output to input (i.e. large conversion efficiency). On the other hand, home retrofits and LED light bulbs can save energy and efficient energy-saving technologies should reduce energy consumption relative to alternative technologies (see column 2 of Table 1). In both cases, the efficiency of technologies is difficult to verify because a counterfactual relative to which efficiency is measured is missing. While the conversion efficiency, may be easy to measure in a laboratory setting, recent evidence of fraudulent manipulation of the conversion efficiency shows scope for deliberate misreporting (see U.S. Environmental Protection Agency, 2017). The energy consumption of a LED light bulb might be easier to measure if consumers have installed a smart-meter.

Putting energy-efficient technologies to use furthermore either demands complex installation steps (such as PV systems and home retrofits) or they can be used as purchased (such as cars, fridges and LED light bulbs). The second domain therefore distinguishes between installation goods (first row of Table 1) and appliances (second row of Table 1). The actors on the supply-side of energy efficient technologies differ along the second domain. Installation goods are typically bought and installed from installers which are in many cases also hired repeatedly to maintain the goods. The quality of such firms ultimately affects the efficiency of technologies to either transform or save energy. In many cases the quality of installation or maintenance is not easy to verify and the effects of reputation on efficiency in markets for credence goods is limited (Dulleck et al., 2011). This implies that installation goods may have a larger credence component, because buyers need to trust installers in receiving an installation of high quality.

By contrast, appliances are typically bought from manufacturers directly or via retail firms. Manufacturers of appliances may have incentives to overstate their efficiency (see Houde, 2018; Goeschl, 2019, who show that energy efficiency labels increase a consumer's the willingness to pay). Furthermore, sellers and retailers might be badly informed about the energy efficiency of their products and thus provide inadequate recommendations to consumers. For example, when car manufacturers overstate the energy efficiency of vehicles, their uninformed retailers may unintentionally recommend a suboptimal vehicle to their consumers. Moreover, even if retailers are

Table 1: The ecological system of energy-efficient technologies

		Energy-transforming	**Energy-saving**
Installation goods	*Examples*	PV systems, heat pumps	Home retrofits
	Supply-side	Installers, high degree of maintenance	Installers, rather high degree of maintenance
	Credence component	high	high
Appliances	*Examples*	Cars, fridges	Lightbulbs
	Supply-side	Manufacturers, sellers, rather low degree of maintenance	Manufacturers, sellers, low degree of maintenance
	Credence component	Potentially low	Potentially low

Notes: This Table illustrates the ecological system of energy-efficient technologies. I categorize technologies along two domains, differentiating between i) their intended use to either transform or save energy and ii) their setup to use which either needs an installation or can be used as purchased. Examples, actors on the supply-side and implications for the credence component are listed in the respective cells.

well informed about energy characteristics of products, they may choose not to disclose it and Milgrom (2008) discusses the role of liability rules for withholding information or even for not providing relevant information which the retailer should have known. In practice, this may be difficult and consumers therefore face the difficult task to find a knowledgeable and trustworthy retailer. However, consumer advocacy groups and also regulators may be able to establish trustworthy labeling in the first place and thereby decrease the the credence component of appliances.

The different implications of verifying energy efficiency, installation-quality and trustworthiness of information imply that the credence component may be most acute for installation goods. For appliances, trustworthy information on energy efficiency may take away the credence component. In contexts where consumers have further calculated that the monetary benefits from investing in an energy-efficient car or LED light bulb will outweigh the investment costs, the investment decision is akin to one for an experience good (where consumers ex ante do not know what they need but observe ex-post whether the technology suits their needs) or even normal good (where consumers know ex-ante which technology they need).

In Chapter I, I lay out the basic theory for this thesis and study how asymmetric information between consumers and expert sellers affects the ecological system for energy-efficient technologies. The theoretic model allows to discuss verifiability and liability, two important factors which can increase market efficiency but have limited impact in markets for energy-efficient technologies.[2] Furthermore, I study four

---

[2] Following Dulleck and Kerschbamer (2006) a liability rule makes expert sellers "liable for the provision of inappropriate low quality" (12). The literature of law and economics provides a more

dimensions which can potentially reduce supply-side inefficiencies i) information, ii) separation of diagnostic and treatment, iii) third party-reimbursements, iv) reputation.

First, I highlight that more informed consumers are likely to receive more consumer-friendly services. At the same time, energy labels and certification schemes are crucial to make an informed decision and third party verification and strict liability rules for labels are important to foster trust. Second, separating diagnostic from treatment can reduce inefficient behavior by expert sellers, but the additional cost of audits can also reduce overall market efficiency. Third, evidence from credence goods markets suggests that subsidizing energy-efficient technologies is likely to increase supply-side inefficiencies. Fourth, providing empirical evidence on realized energy savings and associated financial implications for already completed projects may help expert sellers to build a better reputation.

In Chapter II, I study how third party reimbursements, such as subsidies, trigger second-degree moral hazard and hence increase supply-side inefficiencies in markets for solar photovoltaic systems. For this purpose, I use data from the California Solar Initiative (CSI) which is the largest solar subsidy program in California. Importantly, an inquiry launched by the US Congress alleges that installers of PV systems may have incentives to increase short-term cash flows in order to increase their market value (Salzman, 2013). In the context of the CSI data, exerting second-degree moral hazard with the objective to increase short-term cash flows would affect different outcomes depending on the ownership of the PV system. In the empirical analysis, I therefore consider two domains. First, I account for different ownership because PV systems are either third-party owned (TPO) or home-owned (HO) and the CSI pays subsidies to the owner of systems. Therefore, after installation, TPO installers appropriate the subsidies and consumers pay the transaction price via dispersed leasing rates while HO installers receive the transaction price and consumers appropriate the subsidies. Second, I focus on two different outcome variables because TPO installers following the objective to maximize their short-term cash flows would be expected to increase total subsidies while HO installers would be expected to increase transaction prices. In the empirical analysis, I therefore separate the sample by ownership and use variations of subsidy levels afforded by the design of the CSI program to measure the effect of a one dollar

---

sophisticated discussion of the concept of liability. For example, Rottenberg (1965) argues that liability is in most cases derived from fault which has to be intentional and should undoubtedly be attributed to the actions of the expert seller. In the context of credence goods, it is difficult to attribute fault to the actions of the expert seller because for example a short-coming in energy savings from home retrofits may be attributed to a change the consumer's energy consumption or other external factors.

per Watt increase of the subsidy level on the self-reported expected electricity output determining the total subsidies received and transaction prices. I further make use of an instrumental variable strategy to address potential concerns about self-selection into specific subsidy levels.

The results show evidence suggesting that second-degree moral hazard is highly relevant in the context of upfront subsidies and leads to a behavior in line with the objective of installers to maximize short-term cash flows. First, I find that a one dollar increase of upfront subsidies is associated with a statistically significant 0.5 percentage point increase of self-reported expected electricity output of residential systems owned by third parties. This association is non-existent when systems are subject to a mandatory field inspection, suggesting that increased verification prevents installers from reporting exaggerated system characteristics. I further find a significant marginal effect of the subsidy level of home-owned systems, suggesting that a one dollar increase of the subsidy level increases the transaction prices by 3.5 to 7.5 percent at the sample mean.

In Chapter III, I study how giving small gifts to expert sellers can trigger reciprocal behavior and thereby increase consumer welfare in markets for credence goods. In particular, I setup a laboratory experiment which is motivated by findings from the behavioral economics literature suggesting that gift exchange can increase efficiency in markets with hidden actions (Akerlof, 1982; Falk, 2007). In the theoretic framework of Chapter II, I show that expert sellers with a high disposition for other-regarding preferences are likely to provide better and cheaper services to consumers. Offering a coffee to an installer of a solar photovoltaic system could therefore circumvent the tedious task of finding an expert seller with other-regarding preferences and thus reduce supply-side inefficiencies (Kerschbamer et al., 2017).

In the baseline condition (BASE) I document the behavior of consumers and expert sellers without the possibility to gift. In the gift exchange (GE) treatment, I extend BASE by giving consumers the possibility to gift the expert seller before providing services. The conditional gift exchange (GEC) treatment, is similar to GE but the gift is transferred only if the expert seller supplies a service of sufficient quality. Results from this treatment can also inform barriers to energy-performance contracting, where some part of the payments to expert sellers is conditional on reaching a predefined threshold of energy savings. Based on the reciprocity model of Falk and Fischbacher (2006), I hypothesize that gifting can induce expert sellers with a preference for reciprocity to provide more consumer-friendly services. At the same time, expert sellers are expected to perceive the unconditional gift as more kind as the conditional gift.

Results from the experiment show that consumers who interact with expert sellers gift in 35 percent of interactions in the GE treatment and 45 percent in the GEC treatment. The likelihood of undertreatment (i.e. receiving an insufficient service) declines by about 18 percentage points following an unconditional gift and by around 14 percent after a conditional gift. Moreover, an unconditional gift increases undercharging (i.e. being charged the price for the cheap, while receiving the expensive service) by around 18 percentage points and reduces overcharging (i.e. being charged the price for the expensive, while receiving the cheap service) by around eight percentage points, whereas undercharging and overcharging are not significantly affected by conditional gifts. These results are consistent with the presence of expert sellers with preferences for reciprocity and suggest that a conditional gift is perceived as less kind than an unconditional gift.

This thesis provides several new insights for the literature on credence goods and particularly energy-efficient technologies. The framework presented in the first Chapter shows that research in energy efficiency bears important insights for the wider literature on credence goods, as for example contracting over energy savings/production can be applied to other contexts. For instance, payments to healthcare providers could be spread over time and conditioned on the patients' satisfaction with the treatment. Next, I document that stylized findings from other credence goods markets are relevant for investment in energy-efficient technologies. The results presented in Chapter II are the first to document evidence of second-degree moral hazard in the context of photovoltaic systems and further inform the discussion on how subsidies should be designed in order to maximize their cost-efficiency. Finally, I show that gift exchange may improve market outcomes and further circumvent the difficult task to identify prosocial expert sellers in order to receive more consumer-friendly services (Kerschbamer et al., 2017). Moreover, the findings in Chapter III contribute to the literature on i) bonus payments in principal-agent models showing that gifts have a positive effect even when the agent has no reputational concerns and ii) reciprocity and observability showing that expert sellers reciprocate unconditional gifts with a range of unobservable actions (e.g. abstain from overpricing), while conditional gifts only affect observable actions (such as abstaining from undertreatment).

The next three Chapters of this thesis present the self-contained papers one after the other. The thesis then concludes by summarizing results and relating them to policy implications and suggestions for future research. In particular, I show that aligning the incentives of the supply- and demand-side via energy performance contracting may help to prevent supply-side inefficiencies related to the credence component of energy-efficient technologies. However, evidence of the functionality of energy performance contracting in the field is needed. Furthermore, peer effects may help to foster the adoption of energy-efficient technologies, especially if early adopters provide their peer group with verified and hence credible information. I finally discuss the welfare effects of supply-side inefficiencies related to the subsidization of PV systems and associated second-degree moral hazard. While an extensive quantification of other supply-side inefficiencies is missing, I present recommendations for consumers and also regulators which come at a low cost or even for free.

Chapter I

Asymmetric information on the market for energy efficiency: Insights from the credence goods literature

This chapter is mainly based on a paper co-authored by Bruno Lanz (University of Neuchâtel) and has been published in ***The Energy Journal***.

Lanz, B. and Reins, E. (2021). Asymmetric information on the market for energy efficiency: Insights from the credence goods literature. *The Energy Journal*, 42(4).

Abstract

Asymmetric information is an important barrier to the adoption of energy-efficient technologies. In this paper, we study supply-side implications of the associated incentive structure. We build on existing evidence that, in some settings, energy-efficiency owns a credence component, whereby the supply side of the market has more information about what technology is best for consumers. The literature on credence goods markets suggests that an information advantage by expert sellers leads to market inefficiencies, including low trade volume. We start by developing a simple framework to study supply-side incentives related to the provision of energy efficient technologies. We then document inefficiencies and potential remedies by discussing linkages between an empirical literature on credence goods and that on the market for energy efficiency. Finally, we identify implications for the design of policies promoting the adoption of energy-efficient technologies.

Keywords Energy efficiency, Asymmetric information, Credence goods, Energy policy, Environmental externalities, Technology adoption.
JEL codes D18; D82; H23; Q41.

1 Introduction

Energy is consumed for the services it provides, and consumers need a technology to transform energy into these services. Energy-efficiency measures how much of these valuable services can be obtained for a given unit of energy input. It follows that, by adopting more efficient energy-transforming technologies, consumers can potentially lower energy use without affecting the amount of services they consume.[1] Because of externalities associated with energy use, and in particular fossil resources that contribute to both local (e.g. airborne particulate matter) and global (e.g. carbon dioxide) emissions, many countries actively promote the adoption of energy-efficient technologies in order to reduce energy consumption (Gillingham et al., 2016b). As highlighted by Allcott and Greenstone (2012), these policies ought to target inefficiencies on the market for energy technologies. The existing literature emphasizes informational failures that affect investment behavior, including imperfect information about and inattention to future energy savings (see also Gerarden et al., 2017).

We argue that, for a number of relevant energy-transforming technologies, the supply-side of the market may hold relevant information and have little incentives to share it with consumers. Based on this observation, this paper focuses on supply-side incentives associated with asymmetric information on the market for energy technologies and implied market inefficiencies.[2] As initially put forward by Sorrell (2004) and recently discussed by Giraudet (2020) and Plambeck and Taylor (2019), the type of informational asymmetry characterizing these energy-transforming technologies can be conceptualized under the notion of "credence goods," a class of goods for which consumers have incomplete knowledge about own needs both before and after purchase (Emons, 1997). For these goods, consumers have to trust information by an expert seller who can perform a diagnostic and supply a particular product. In the case of energy-transforming technologies, consumers typically do not directly observe the level of energy-efficiency of alternative technologies, and may not be able to ascertain which technology is cost-minimizing, both before and after purchase/installation.[3] In turn,

---

[1] Note that energy-efficiency improvements reduce the relative price of energy services, which may lead to an increase in the demand for these services. Therefore, improving energy-efficiency does not imply a one-to-one reduction of energy consumption (see Chan and Gillingham, 2015).

[2] Energy-transforming technologies may require an installation (e.g. heating systems) or can directly be used by consumers (e.g. technologies such as a fridge or a car). As we discuss below, the quality of both the technology and of the installation affect realized energy savings and can give rise to a component of trust in information provided by the supply-side of the market.

[3] Many energy-transforming technologies are subject to governmental labeling regulations, which inform consumers about energy-efficiency and mitigate information asymmetry. However, using this

when the supply-side of the market possesses an informational advantage, decisions to invest in energy-transforming technologies inherit the properties and inefficiencies identified in markets for credence goods.

In Section 2 we start by considering energy-efficiency investment decisions in relation to the basic credence goods model of Dulleck and Kerschbamer (2006). This framework allows us to identify sources of inefficiencies associated with credence goods, and discuss “baseline” results from the seminal implementation of credence goods markets in the laboratory by Dulleck et al. (2011). In the next step, we relate the framework of Dulleck and Kerschbamer (2006) to the simple model for energy-efficiency investments by Allcott and Greenstone (2012), clarifying which aspects of investment decisions are likely to be affected by informational asymmetries. This delivers the main contribution of this work, namely identifying how and when inefficiencies studied in the credence goods literature translate in the context of energy-efficiency investment decisions, and how the credence component of energy-using technologies can affect market efficiency.

Conceptually, the credence goods framework provides a supply-side perspective on the observed tendency of consumers to invest “too little” in energy-efficiency, seemingly failing to realize financial benefits from energy savings (the energy-efficiency gap, see Jaffe and Stavins, 1994a). Indeed, theoretical studies on credence goods such as Emons (1997) and Dulleck and Kerschbamer (2006) suggest that asymmetric information induces a reduction in trade volume on credence goods markets. The necessity to trust expert sellers comes from the possibility of inefficient supply-side behavior, which can be classified under three possible outcomes: (i) expert sellers supply a lower quality than what the consumer needs (undertreatment), (ii) the quality supplied is higher than what is needed (overtreatment), and (iii) expert sellers charge for goods or services that are of higher quality than what is actually supplied (overpricing). While asymmetric information is only one of the factors affecting the energy-efficiency gap, and may not affect all consumers and technologies equally, understanding how the market fails in relation to supply-side incentives is important for the design of public policies.[4]

---

information to determine a cost-minimizing option can remain challenging for some consumers (e.g. Brounen et al., 2013). Moreover, as we discuss extensively below, a lack of ex-post verifiability implies supply-side incentives to manipulate information provided to consumers, even in the presence of labels (Houde, 2018; Goeschl, 2019).

[4] Gerarden et al. (2017) distinguish between three categories of factors explaining an energy-efficiency gap: market failures (including asymmetric information), deviations from the canonical behavioral framework (e.g. loss aversion), and flaws in the modeling framework such as incorrect cost calculations for some of the options. As we discuss in the text below, some behavioral and modeling flaws can be seen as originating from the credence component of energy technologies.

Dulleck and Kerschbamer (2006) suggest two institutional features of credence goods markets that can potentially restore market efficiency without external intervention. First, *verifiability* refers to a case in which consumers are able to verify the characteristics of the product after purchase/installation. Second, *liability* represents a case in which expert sellers are liable to solve the consumers' problem. Under specific conditions, which we discuss below, either verifiability or liability leads to efficient trade in markets for credence goods. For many energy-efficiency investments, however, neither verifiability nor liability are likely to solve the information problem. The key reason is the difficulty (or costliness) for consumers to ascertain, for each possible technology available on the market, actual energy savings that will be achieved. Also ex-post, after purchase and/or installation, measuring energy use per unit of service and defining a valid counterfactual remain a challenging endeavor (see e.g. Joskow and Marron, 1992; Burlig et al., 2020). Realized energy savings are influenced by exogenous factors (e.g. the weather) and endogenous factors (e.g. changes in the demand for energy services). This leads us to discuss energy performance contracting as a way to align supply- and demand-side interests (see e.g. Sorrell, 2007; Klinke, 2018). However, insights from the credence goods literature suggest that the difficulty to credibly quantify energy savings for many relevant energy-transforming technologies limits the range of application of performance contracting.

In Section 3, we turn to a review of the empirical evidence from the credence goods literature, building on the work of Kerschbamer and Sutter (2017). This part of the paper overviews results from experimental markets for credence goods, as well as field evidence for products such as car repairs, taxi rides, and medical treatments.[5] Our objective is to use the drivers of market inefficiencies identified in the credence goods literature to organize and discuss selected contributions to the energy-efficiency literature. This affords the second main contribution of our paper, which is to provide a rejoinder of two separate but related streams of the economics literature, and thereby offer a novel perspective on supply-side incentives underlying investment decisions in energy-transforming technologies.[6]

---

[5] In our view, laboratory experiments and field studies provide complementary evidence. The former allows to systematically vary selected institutional aspects of decisions in a controlled environment, while the latter quantifies the magnitude of inefficiencies for real-world decisions.

[6] We note that empirical evidence derived from other credence goods markets may not apply directly in the context of energy-efficiency investments. We argue, however, that many credence goods markets display patterns that are consistent with predictions from the model by Dulleck and Kerschbamer (2006), and that includes observations resulting from the energy-efficiency literature. Linking empirical results from different domains is therefore important.

Concretely, we consider four important characteristics that affect supply-side behavior in markets for credence goods. First, we discuss the degree of informational asymmetry between consumers and expert sellers, and the conditions under which mitigating the information gap reduces market inefficiency. Experimental evidence by Balafoutas et al. (2013) suggests that consumers who signal to be informed about the characteristics of a credence good are more likely to receive a correct treatment. In the context of energy-efficiency, informing consumers is directly related to the use of energy efficiency labels and certification (see e.g. Newell and Siikamäki, 2014). Our reading of the literature leads us to emphasize the trust component in energy-efficiency labels, and the role for independent third parties (or competing experts) to test whether the actual energy intensity of technologies corresponds to the certified energy efficiency. Moreover, in an effort to enforce trust on the market for energy-efficiency, we highlight the need to make suppliers liable to deliver products that are in line with the certification.

The second dimension of credence goods markets we consider is the necessity to carry out a diagnostic, and the ensuing possibility to separate diagnostic from treatment. In the context of energy efficiency, this can take the form of independent energy audits. Supply-driven inefficiencies are expected to decrease if one expert is paid to perform the diagnostic and another expert is paid to perform the treatment, as the diagnosing expert has no incentives to recommend inappropriate products. The literature on credence goods, however, suggests that separate diagnostic can also worsen inefficiencies because it introduces an additional cost to market participation (Greiner et al., 2017; Mimra et al., 2016b). We further discuss the possibility of ex-post auditing (Allcott and Greenstone, 2017) as another approach to improve market efficiency in a credence goods framework. Policies that lower the cost of third party audits, both before and after purchase/installation can mitigate the credence component of energy technologies and favor trust in the behavior of expert sellers.

Third, we discuss how third party reimbursement reduces market efficiency in credence goods markets (e.g. Kerschbamer et al., 2016; Balafoutas et al., 2017). In particular, empirical evidence shows that experts are inclined to overtreat and overprice consumers whenever a third party (e.g. an insurance) covers the cost of the treatment. The use of subsidies for energy-efficient technologies makes these results highly relevant for the design of energy and environmental policy. We relate results from credence goods markets to empirical evidence on consumer subsidies and sales incentives studied by Allcott and Sweeney (2017). Overall, this suggests that the presence of asymmetric information diminishes the effectiveness of subsidies in the adoption of energy-efficient technologies, although further research on how subsidies affect pricing behavior by expert sellers is needed.

Finally, we examine the role of reputation and repeated interactions in the context of credence goods in general and energy-efficiency in particular. The basic credence goods model suggests that honest expert sellers will be driven out of the market, which is reminiscent of the lemons problem discussed in Akerlof (1970). Providing mechanisms for expert sellers to signal their trustworthiness, such as neutral third parties publishing credible information about quality of service received, can contribute to help expert sellers establish a good reputation. We emphasize, however, that the difficulty to quantify energy savings is again crucial for reputation-building by expert sellers and whether consumers trust reputation information (see Gillingham and Tsvetanov, 2018).

The paper concludes in Section 4 by summarizing policy implications of our work and bringing together some suggestions for further research.

2 Credence goods and energy-efficiency

This section discusses the relationship between credence goods and energy-efficiency. First, we briefly describe information asymmetries associated with credence goods and implied market inefficiencies using the general framework of Dulleck and Kerschbamer (2006) and the related experimental procedure by Dulleck et al. (2011). Second, we present a simple representation of decisions to invest in energy efficiency and clarify the sources of asymmetric information that may affect specific technologies and consumers on the market for energy efficiency. Finally, we discuss verifiability and liability in the context of energy-efficiency investments, as well as the role of energy performance contracting.

2.1 Credence goods and market inefficiencies

As in Dulleck and Kerschbamer (2006), consider a consumer with a problem that can be either minor or severe. The consumer, however, does not know which of the two conditions he faces, and hence whether an expensive treatment q_h or a cheap treatment q_l is needed. This classification mirrors a setting in medical treatments or car repairs, but it can also be interpreted more broadly as representing preferences of a consumer for a particular product or service.[7] To assess which good q_h or q_l is needed, the consumer

---

[7] As discussed in Dulleck and Kerschbamer (2006), the framework can be adapted to a situation in which consumers' type is drawn from a continuous distribution, and they face multiple differentiated products. For the present discussion, however, a binary case is sufficient to identify the nature of

relies on an expert seller to perform a diagnostic. The expert seller then recommends either q_h or q_l, supplies the good, and charges either p_h or p_l (with $p_h \geq p_l$). Note that, since the consumer does not observe his condition after treatment, either price can be charged regardless of the good supplied.

This simple setting is the basis of the experimental market for credence goods studied in Dulleck et al. (2011) and implemented as a stage game in which one consumer interacts with one expert seller. Throughout the game, each consumer only knows that q_h is needed with probability h and q_l with probability $1 - h$ (h is set to 0.5). The expert seller faces a cost for performing high vs. low treatments, with $c_h > c_l$, and these costs are common knowledge. One implication is that, by observing prices, the consumer can determine markups and hence incentives for expert sellers to supply either good.

In the first stage, the expert seller posts prices p_h and p_l, and the consumer decides whether he wants to participate in the market or not. If the consumer opts out, the stage game stops and both participants receive an outside option $o > 0$. If the consumer opts in, the game moves on to a second stage in which the expert seller learns about the severity of the problem faced by the consumer (akin to a diagnostic), elects to supply either q_h or q_l, and charges either p_h or p_l (independently of the treatment supplied). At the end of the stage game, the payoff for expert sellers is the difference between the price charged and the cost of the good supplied: $\pi_s = p_{\text{charged}} - c_{\text{supplied}}$. For the consumer, if the problem is solved ($q_{\text{supplied}} \geq q_{\text{needed}}$), the payoff is $\pi_c = v - p_{\text{charged}}$, where $v > 0$. If instead $q_{\text{supplied}} < q_{\text{needed}}$, $v = 0$ and the consumer gets $\pi_c = -p_{\text{charged}}$.

In this setting, market efficiency requires that the sum of surpluses is maximized. With the baseline parametrization, this occurs when the consumer opts into the market, and the expert sellers recommends the appropriate treatment. However, asymmetric information gives rise to three types of supply-side behavior that lower trade and reduce market efficiency. First, *undertreatment* occurs if the consumer needs q_h, but the expert seller supplies q_l. This implies that the problem of the consumer is not solved, hence $v = 0$. Second, *overtreatment* occurs when the consumer needs q_l but receives q_h. In this case the problem is solved ($v > 0$), but some of the tasks performed by the expert seller are unnecessary. Third, *overpricing* is a situation in which the consumer receives q_l and is charged p_h, so that the consumer pays for a quality that he did not receive.

A purely selfish expert seller who maximizes own surplus always supplies q_l and charges p_h (i.e. $\pi_s = p_h - c_l$). For consumers who need q_l this implies overpricing, while for those

---

the credence goods problem, and is also in line with the model by Allcott and Greenstone (2012) discussed below.

needing q_h it implies undertreatment. Experimental results by Dulleck et al. (2011) show that 88 percent of consumers in the q_l condition are subject to overpricing, and 53 percent of consumers in the q_h condition are undertreated. By contrast, overtreatment is only observed in 6 percent of all interactions, as it is always dominated by overpricing $(p_h - c_l > p_h - c_h)$.[8]

Similar to the lemons problem in Akerlof (1970), profit maximizing expert sellers who undertreat and overprice drive out of the market those who install adequate quality and charge adequate prices. Moreover, consumers who expect selfish behavior by expert sellers are better off opting out of the market ($\pi_c = o$), which in equilibrium leads to market collapse. In Dulleck et al. (2011), the share of consumers who opt out increases from around 40 percent in the first period to about 80 percent in the last period. Furthermore, consumers who are undertreated in period $t-1$ are significantly more likely to opt out of the market in period t. While complete market breakdown does not occur, mainly because some expert sellers display other-regarding preferences and do not undertreat (see also Kerschbamer et al., 2017), the low level of trades is associated with significant market inefficiencies.[9]

2.2 Energy-efficiency as a credence good: A simple framework

We now turn to energy-transforming technologies, which include heating and/or cooling systems, vehicles, lighting equipments, and electric appliances such as fridges or freezers. For all these products, alternatives available on the market are typically differentiated with respect to their energy-efficiency (among other things). In the following, we discuss the credence nature of energy-efficient technologies in the context of a simple model, combining the representation of energy-efficiency investment decisions by Allcott and Greenstone (2012) with the primitives of credence goods model by Dulleck and Kerschbamer (2006) discussed above.

Consider two different versions of an energy-transforming technology, namely an energy-efficient version q_h and an energy inefficient version q_l. The upfront price for each option p_h and p_l covers the technology and its installation, and subsequently

---

[8] This is a result of the parametrization of the experiment. In real world markets for credence goods, overtreatment is frequently observed (see e.g. Rasch and Waibel, 2017; Baniamin and Jamil, 2018; Gottschalk et al., 2020). Note also that the results from this experimental procedure were successfully replicated by Camerer et al. (2016).

[9] Beck et al. (2014) replicate this experimental design with car mechanics instead of a students subject pool, with comparable results. A key difference is that mechanics are found to be more likely to overtreat, although the difference declines as the game is repeated.

consumers pay the energy cost associated with the quantity of energy services consumed, denoted m_i.[10] As per Dulleck et al. (2011), consumers can be of either one of two types, a high usage type m_h and a low usage type m_l, with $m_h > m_l$. The type of consumers in turn determines the cost-minimizing technology: for m_h the more efficient version q_h is cost minimizing, whereas for m_l technology q_l is cost minimizing.

Formally, q_h is cost-minimizing for consumer i whenever the following inequality holds:

$$\frac{p_e\, m_i\, (e_l - e_h)}{1+r} \quad > \quad p_h - p_l\,, \tag{1}$$

where p_e is the private unit cost of energy, e_l and e_h represent energy intensity of each technology (with $e_l > e_h$), and $r > 0$ is a risk adjusted discount rate.[11] We argue that, in many instances, the consumer cannot know whether the inequality holds, and hence whether he needs q_h or q_l. First, energy-efficiency as measured by e_l and e_h cannot be directly observed by consumers. Instead, consumers need to rely on external information in the form of engineering or sales agent expertise (unless there is some certification, something we discuss below). Second, while some consumers may be relatively well informed about their usage rate m_i, it is typically imperfectly observed, and there is potentially some cost to discover it. Consumers may also lack the expertise to translate information about m_i into energy use (e.g. in kWh), be unaware of energy prices p_e, or may have difficulties to perform present-value comparisons. By contrast, an expert seller can ascertain both e_l and e_h, and determine the type of consumers m_i for whom inequality (1) holds.

A crucial component of this framework are energy savings ($p_e m_i(e_l - e_h)$), which determine the cost-minimizing technology, and asymmetric information on its constituting elements can give rise to market inefficiencies and low trade volume inherent to credence goods (Sorrell, 2004; Giraudet, 2020). In particular, while we discuss evidence on supply-side inefficiencies in detail below, possible outcomes identified in the credence goods literature can be summarized as follows. First, undertreatment occurs when q_h is cost-minimizing (i.e. the consumer is of type m_h), but instead q_l is provided (independently from the price charged).[12] Second, overtreatment

---

[10] The framework can easily accommodate a situation in which the inefficient option is the status quo, as in Allcott and Greenstone (2012), by setting $p_l = 0$. Considering instead a replacement decision is closer to the setting of Dulleck et al. (2011) in which the consumer faces a problem and seeks a solution.

[11] For simplicity, usage-rate m_i is assumed to be unaffected by the type of technology installed. This abstracts from a rebound effect and is in line with Allcott and Greenstone (2012).

[12] Intuitively, for high usage m_h consumers, savings on energy expenditures associated with the energy-

implies that q_l is cost-minimizing for the consumer (i.e. the consumer is of type m_l), but the expert seller delivers q_h (independently from the price charged). Third, overpricing corresponds to a case in which q_l is installed but the price of q_h is charged.

Importantly, in practice there is heterogeneity across both consumers and energy technologies. For example, a consumer on the market for cars may know average distance driven per year (m_i), the price of gas (p_e), comprehend fuel-efficiency information provided by car manufacturers (e_l and e_h), and be able to identify a cost minimizing alternative. Empirical evidence for the general population, however, suggests that energy literacy is relatively low. For example, using a survey of Dutch households, Brounen et al. (2013) reports that only 56% know their energy expenditures and 40% were not able to make a comparison between long-term savings and upfront investment cost (akin to equation 1; see also Jessoe and Rapson, 2014 and Blasch et al., 2019 for further evidence on energy literacy). Moreover, supply-side profit maximization motives imply incentives to manipulate information, which may lead to fraudulent behavior even in the presence of sophisticated reporting standards such as those applied in the U.S. car industry (see U.S. Environmental Protection Agency, 2018). For other energy-efficiency investments, such as thermal insulation, supply-side (e.g. architects) informational advantage seems even more acute.

Ideally, consumers on the market for energy-transforming technologies would perfectly observe realized energy savings for all technologies, and based on this determine which technology is best suited for their needs. When this cannot be observed, investment decisions involve a component of trust (independently from perceived risks, see Gillingham and Palmer, 2014). In turn, this may lead some consumers to stay out of the market, thereby contributing to the energy efficiency gap. Based on this, we proceed with a discussion of how properties of credence goods markets contribute to the extent of market inefficiencies, and relate these to energy-efficiency investment decisions and policies affecting these.

2.3 Verifiability, liability, and energy contracting

Dulleck and Kerschbamer (2006) highlight two features of credence good markets that can restore market efficiency without external intervention, namely verifiability and liability. In the following, we discuss both institutional features in the context of energy-efficient technologies. This leads us to consider energy performance contracting as a

---

efficiency level e_h more than compensate higher investment costs p_h.

response to the credence component of energy-efficiency investments.

Verifiability and energy efficiency

Under verifiability, consumers are able to identify, after treatment, whether the expert seller has installed q_h or q_l. As a consequence, verifiability rules out the possibility of overpricing, and the expert will supply the treatment that maximizes his profits. In the context of Dulleck and Kerschbamer (2006), where consumers know c_h and c_l, expert sellers can only attract consumers under equal markups: $p_h - c_h = p_l - c_l$.[13] Together with an assumption that expert sellers install the appropriate quality whenever they are indifferent, this leads to the efficient outcome (see Emons, 1997, for a similar result).

Experimental results from Dulleck et al. (2011) indicate, however, that verifiability does not increase market efficiency compared to the baseline. There are two main reasons for this. First, equal markup prices are posted in only four percent of interactions, which gives rise to incentives to overtreat or undertreat. Second, some expert sellers display antisocial preferences, leading them to supply the inappropriate treatment even if it generates lower profits (see Kerschbamer et al., 2017, for further evidence on this). The impact of verifiability on market efficiency in experimental credence goods markets is therefore limited.

In the context of energy-efficiency, verifiability requires that consumers are able to observe whether technology q_h or q_l is operating, which implies measuring (verifying) realized energy savings $\hat{m}_i(\hat{e}_l - \hat{e}_h)$ after purchase/installation. For a number of energy-transforming technologies, it may be possible to monitor energy use (e.g. by measuring it with a smart meter), although this requires time and effort. Moreover, one empirical challenge to verifiability is the definition of a valid counterfactual for energy consumption (e.g. with an inefficient technology, see Joskow and Marron, 1992). In fact, a branch of the literature on energy efficiency is dedicated to the estimation of energy savings achieved by improving energy-efficiency in buildings. For example, Dubin et al. (1986) and Fowlie et al. (2018) use randomized controlled trials to estimate counterfactual energy consumption of renovated buildings, while Burlig et al. (2020) employ machine learning techniques to predict energy use in the absence of energy retrofits. However, for individual investors the difficulty to quantify realized energy savings remains.

---

[13] If the markup associated with one of the treatments is higher, the associated product can be expected to be supplied independently of the actual condition faced by the consumer.

An absence of verifiability creates incentives for the supply-side of the market to manipulate information about expected energy savings. Therefore, the framework suggests that investments in energy-efficient technologies are prone to a systematic bias of ex-ante projections for energy savings $m_i(e_l - e_h)$ relative to realized energy savings measured after purchase/installation $\hat{m}_i(\hat{e}_l - \hat{e}_h)$. This is in line with early empirical evidence from Nadel and Keating (1991), which reports that engineering estimates of savings tend to be higher than empirical measurements. Similarly, Davis et al. (2014) estimate energy savings from a program to replace inefficient refrigerators in Mexico, finding that these amount to about one quarter of predicted savings. The above-cited studies on buildings retrofits also confirm that realized energy savings tend to be systematically below ex-ante estimates.

In sum, for many energy-transforming technologies, verifiability is both difficult and costly, which can explain differences in measures of energy savings before and after purchase/installation. In turn, the credence goods model suggests that this may lead consumers to stay out of the market, reducing investments in energy-efficiency. From a policy perspective, this may also induce policy-makers to incentivize behavior that is not cost-effective (Fowlie et al., 2018).

Liability and contracting for energy-efficiency

Under liability the expert is made liable for supplying an appropriate treatment. In the setup of Dulleck and Kerschbamer (2006) and Dulleck et al. (2011) a liable expert does not have the possibility to install q_l if the consumer needs q_h, ruling out undertreatment. Experimental results suggest that the associated reduction in the expected loss by consumers nearly doubles market participation, as measured by the number of trades. Market efficiency is, however, hampered by overpricing, which is observed in 75 percent of trades.

While liability can easily be introduced in a laboratory environment, in the field it may not be feasible to rule out undertreatment (e.g. through fines). In the case of energy-efficiency, quality control is difficult to enforce (Gerarden et al., 2017), incentivizing expert sellers to cut down costs at the installation stage and undertreat consumers. For example, Giraudet et al. (2018) provide evidence that energy savings achieved by attic insulation and duct sealing interventions are lower when installed on a Friday, and interpret this as a change in labor costs before the weekend. Despite this realization of undertreatment, it is difficult to enforce penalties for relatively low energy savings, as energy savings are affected by energy consumption which may fluctuate for exogenous reasons.

One specificity of energy-transforming technologies is that they are often sold by retailers, which act as intermediaries between manufacturers and consumers. Retail agents may not be well informed about the energy-efficiency of products (e_l and e_h), or be able to carry-out a diagnostic to determine the most appropriate technology. In this setting, intermediaries who cannot be made liable for false statements engage in cheap talk which can hardly influence the buyer, and may reinforce the asymmetric information problem (see Farrell and Rabin, 1996, for a discussion). Moreover, even if intermediaries are well informed about energy characteristics of products, they may choose not to disclose it to a majority of uninterested consumers (Allcott and Sweeney, 2017). As a solution, Milgrom (2008) discusses the role of liability rules for withholding information or even for not providing relevant information which the sales agent should have known, although in practice this may be difficult.

One approach to enforce liability is energy performance contracting, whereby a contractor (manufacturer or retailer) selects and installs a technology, and is subsequently entitled to a share of the financial gains associated with realized energy savings. The structure of the contract should be such that expert sellers maximize profits by installing the cost minimizing technology (Tietenberg, 2009). Energy performance contracts would therefore make undertreatment unattractive, as expert sellers would forgo profits associated with energy savings. Similarly, installation costs are borne by expert sellers, which prevents overpricing. The contract may also contain incentives for expert sellers to maintain or improve energy savings (see Sorrell, 2007).

In the context of equation (1), a performance contract can be defined as an insurance on realized energy savings, whose financial value $p_e(\hat{m}_i\hat{e}_l - \overline{m}_i\overline{e})$ is shared by contract holders. However, in practice, the difficulty to define baseline energy consumption $\overline{m}_i\overline{e}$ is an important barrier to performance contracting. While this has led the contracting industry to design standardized estimation procedure, essentially a before/after comparison with ad-hoc adjustments for the consumption of energy services (see Efficiency Valuation Organization, 2018), it can only be applied to specific investments. Second, Klinke (2018) provides survey evidence that economic viability is a significant barrier, mainly because of the risk associated with future energy savings. For example, external factors such as a changing climate may affect realized energy savings and induce a risk that the supply-side may not be willing to hold.

We conclude that liability can improve market efficiency for a limited number of energy-transforming technologies, either through regulation and fines or through energy performance contracting. However, in many cases, exogenous and endogenous variability in realized energy savings hampers the possibility to make expert sellers liable for failure to deliver energy savings in line with ex-ante projections.

3 Empirical evidence from credence goods markets: Implications for energy-efficiency policies

In this section we review existing empirical evidence on credence goods markets and link it to the literature on energy-efficiency. The structure of our argument broadly follows Kerschbamer and Sutter (2017), and we focus on four institutional features that are relevant for energy-efficiency: (i) the degree of information asymmetry and the role of certification; (ii) separation between diagnostic and treatment in relation to energy audits; (iii) third party reimbursement and subsidies to energy-efficient technologies; and (iv) reputation and repeated interactions in a market for emerging technologies.

3.1 Informing consumers and certification

Market inefficiencies associated with credence goods stem from an informational advantage held by the supply-side of the market. Therefore, a natural approach to improve market efficiency is to inform consumers. Indeed, as initially put forward by Darby and Karni (1973), market inefficiencies are proportional to differences in information. Imperfect information on the market for energy-efficient technologies has been identified as an important driver of market inefficiencies (e.g. Allcott and Wozny, 2014; Jacobsen, 2015). For consumers, trustworthy and verifiable information on expected savings $m_i(e_l - e_h)$ (from energy labels or any other sources) would transform energy-efficiency in a search good, thus reducing a key source of market inefficiencies.

Field experimental evidence in the context of taxi rides reported by Balafoutas et al. (2013) quantifies how the supply-side of the market exploits the degree of asymmetric information. The authors find that taxi drivers (expert sellers of cab rides) are more likely to overtreat consumers who explicitly state that they are unfamiliar with the city by taking them on a detour. Moreover, when a consumer signals to be a foreigner, the probability that the driver applies a false tariff and charges extra fees increases, and these are both instances of overpricing. By contrast, consumers who use their smartphone to suggest directions to the driver, thereby signaling some degree of expertise, are less likely to be subject to overtreatment and overpricing.[14]

Labels for energy-using technologies inform consumers about energy consumption and have been introduced in many countries (e.g. U.S. Federal Trade Commission, 1979; European Commission, 2013, for the United States and the EU respectively). Empirical evidence suggests that the information transmitted on labels helps consumers to make cost efficient decisions. For example, Newell and Siikamäki (2014) test alternative designs for information contained in energy labels and emphasize the importance of information on financial savings. A closely related research by Davis and Metcalf (2016) shows that providing tailored information on usage rates (m_i) results in more cost efficient choices of air conditioners.

However, supply-side incentives associated with the credence nature of energy-efficiency may compromise trustworthiness of labels. Goeschl (2019) finds systematic discrepancies between self-declared e^* and verified energy-efficiency ratings of refrigerators sold in the EU market ($\hat{e} - e^* > 0$), which suggest that labels are not sufficient to ensure that consumers use it as a trustworthy source of information. Following the credence goods framework, this can be related to evidence that consumers remain at least partly inattentive to this information (Sallee, 2014; Allcott and Knittel, 2019).

Trust in certification can potentially be addressed by competitor testing, whereby competing expert sellers verify information about products of competitors (i.e. enforce liability for $\hat{e} = e = e^*$). Plambeck and Taylor (2019) show that when violations of certification lead to fines or restricted market access, competitor testing can be more effective in enforcing certification as compared to a regulator. However, entry of non-compliant firms with low-quality products is still possible when the market share of these low quality products is not sufficient to draw competitors to test the products.

Therefore, even if competitor testing is allowed, the need for a regulator to punish violation of product certification remains. Related to this, a regulator may also introduce minimum efficiency standards $\underline{e}$ in order to avoid low-quality products entering the market. Brucal and Roberts (2019) and Houde and Spurlock (2015) analyze a change on minimum energy-efficiency standards for several technologies, reporting evidence that prices of the remaining products decline, while quality and consumer welfare

---

[14] Similarly, in the medical domain, Domenighetti et al. (1993) study the frequency of common surgical procedures and show that the probability to receive a surgery is significantly lower if the patient has a physician in his family or is a physician himself, which presumably reflects lower informational asymmetry (see also Gruber and Owings, 1996; Gruber et al., 1999).

increase. These results suggest that removing highly inefficient technologies from the market can increase overall consumer welfare.

Certification in the context of credence goods may induce expert sellers to increase markups associated with energy-efficient products ($p_h - c_h$). Houde (2018) studies pricing behavior for suppliers of refrigerators who have lost their energy certification, and estimates that certification increases the price of products by 2 to 5 percent.[15] This suggests that certification on markets for credence goods can lead expert sellers to partially appropriate expected benefits associated with lower energy consumption.[16]

Finally, we note that information may also come in the form of general recommendations such as web-based guidelines provided by the U.S. Department of Energy on how consumers can perform energy audits themselves or estimate their energy use for certain services.[17] External information on whether equation (1) is likely to hold may reduce the degree of asymmetric information when a diagnostic is performed and can be used to signal knowledge about personal needs. However, Gottschalk et al. (2020) report field evidence in the context of dental care that such information does not necessarily increases market efficiency. In particular, the authors find that patients who signal that they will obtain information on diagnostics from an internet platform do not benefit from a lower probability of overtreatment.

3.2 Separating diagnostic from treatment and independent energy audits

Energy audits provide information on the appropriate (cost-minimizing) treatment for consumers. However, if it is performed by an expert who also supplies the treatment to the consumer, supply-side incentives may cause a problem of supply-induced demand (e.g. for whichever option affords higher markups).[18] This issue can be addressed by

---

[15] A related paper by Fisher (2010) shows that vehicle manufacturers strategically select fuel efficiency of vehicles to extract surplus from consumers with alternative tastes for this characteristic of vehicles.

[16] Note that when costs are unobserved, consumers are unable to determine whether higher prices reflect higher cost or surplus appropriation by the expert seller. This possibility is ruled out in the setup studied in Dulleck et al. (2011).

[17] See for example the Energy Saver Program of the U.S. Department of Energy (2018). By contrast, the National Energy Audit Tool (U.S. Department of Energy, 2012) is designed for experts on the supply side (utility companies, residential energy professionals, auditors, energy consultants and analysts), rather than for untaught consumers.

[18] Similarly, Causholli et al. (2013) and Knechel (2013) discuss how auditing services (e.g. in the context of accounting) share the characteristics of credence goods when the audit and the treatment is performed by the same agent. They show that auditors may have incentives to under- or over-

separating the diagnostic from the treatment, so that the expert performs the diagnostic (equation 1) while a seller supplies a treatment (q_h or q_l in our two goods framework). While this reduces the scope to exploit asymmetric information, the consumer may have to pay for a diagnostic separately. This creates an additional barrier to the provision of a credence good.

Greiner et al. (2017) studies the separation of diagnostic and treatment experimentally in the context of a physician-patient relationship. When the diagnostic is free, consumers are more likely to seek one and treatment take-up increases. However, even though the seller is forced to stick to the diagnostic provided by the expert, overtreatment still occurs in 20 percent of all transactions (51 percent in the baseline), and undertreatment increases from 7 to 24 percent.[19] These results presumably reflect spiteful behavior by experts since they earn no diagnosis fee. However, when the diagnostic is costly, fewer patients seek a diagnostic and market efficiency declines. These effects are confirmed in an experiment by Mimra et al. (2016b), which shows that the possibility for consumers to obtain multiple diagnostics before they interact with an expert seller lowers the probability of overtreatment, whereas diagnostic fees again reduce overall welfare.

In the context of energy-transforming technologies, Anderson and Newell (2004) study a government-sponsored independent audit program, showing that about half of recommended energy-efficiency measures are adopted, and Fleiter et al. (2012) provides survey evidence that "consultants neutrality" is an important driver of participation in subsidized energy audits. Blonz (2019) evaluates a program to replace refrigerators among low-income households in California, finding that contractors limited to perform a diagnostic (i.e. whether the fridge should be replaced or not) are less likely to recommend replacement as compared to contractors who were further paid to replace the fridge. This can be interpreted as evidence that separating diagnostic from treatment decreases overtreatment in the context of energy-efficiency investments. However, empirical evidence reported in Fowlie et al. (2015) and Fowlie et al. (2018) also confirms that households' welfare cost associated with energy audits is significant, with few households signing up for highly subsidized energy audits. In turn, as suggested by results from the credence goods literature, the overall welfare impact of audits is likely mixed (see also Abrahamse et al., 2005).

---

audit.

[19] In this experiment, consumers can verify which treatment was provided. In combination with the parameters of the experiment, this implies that both consumers and expert sellers benefit more from solving the high condition, so that overtreatment is expected.

A related intervention is that of ex-post auditing, discussed in the field study by Allcott and Greenstone (2017). In a first step, an independent state-certified auditor performs a free diagnostic to assess the consumer situation (m_i) and recommends the level of energy-efficiency (e_h or e_l). In a second step, a certified contractor installs the technology (insulation, heating, cooling). After the work is completed the independent contractor returns to verify that the technology has been installed adequately ($\hat{e}_h$ or $\hat{e}_l$, respectively). Allcott and Greenstone (2017) find that the sequence of audits increases the willingness to pay for unobserved (non-monetary) benefits associated with energy-efficiency, leading 20 to 50 percent of consumers to install a technology with negative financial returns. A combination of ex-ante and ex-post auditing may therefore act as a safeguard to supply-side inefficiencies arising from credence component of energy technology, although the cost of these audits may significantly increase the cost to energy-efficient technology adoption.

3.3 Third party reimbursement, subsidies and markups

Third party reimbursement represents a situation in which an expert seller knows that the price charged will at least partly be borne by a third party, such as an insurer or an employer. Both theoretical predictions and empirical evidence shows that this leads expert sellers to overtreat and overprice consumers, and increase market inefficiencies. We argue that this line of research is relevant for the design of subsidies for energy-efficient technologies (s_h), which often play an important role in policy promoting energy efficiency investments. In particular, subsidizing credence goods is likely to affect pricing behavior or induce subsidy manipulation, and in turn the incentives of an expert seller to provide adequate services.

Kerschbamer et al. (2016) provides field evidence on third party reimbursement in the market for computer repairs. When an expert seller knows that the consumer is insured, the average bill for a pre-specified problem is EUR 129, as compared to EUR 70 when the consumer bears the full cost of the reparation. About one third of the difference is due to overtreatment (performing unnecessary repairs), the rest being explained by overpricing (charging for services which were not provided). Similar results for taxi rides are reported in Balafoutas et al. (2017), as passengers who state that their expenses are reimbursed are more likely to be subject to overpricing. In the healthcare context, field evidence shows that physicians are more likely to overtreat when patients are insured (Iizuka, 2007, 2012; Lu, 2014), and Huck et al. (2016) replicate this finding in a laboratory experiment.[20]

In our framework, third-party reimbursement is akin to a situation where part of p_h and p_l is not borne by the consumer. By contrast, subsidizing the energy-efficient technology implies that consumers only bear $p_h - s_h$, making q_h cost-effective for additional consumers with a lower m_i (see related discussion of *additionality* in Globus-Harris, 2020, and Gilbert et al., 2019). Therefore the traditional view is that subsidizing energy-efficient technologies accelerates their adoption (Jaffe and Stavins, 1994b; Comstock and Boedecker, 2011). However, results from the credence goods literature indicate that expert sellers are prone to manipulate both diagnostic and delivery to appropriate (part of) the subsidy. For example, Iizuka (2007) shows that physicians who can directly sell drugs to patients are more likely to prescribe those that afford higher markups.

Subsidies may therefore affect supply-side incentives, notably through markups. While evidence on how subsidies affect market prices for energy-efficient technologies is scarce,[21] Allcott and Sweeney (2017) provide experimentally controlled evidence on how subsidies and markups affect expert sellers behavior and the demand for water heaters. They study the behavior of sales agents in a call center who cannot manipulate prices and instead rely on sales incentives, finding a strong complementarity between financial incentives for the seller (premium for selling q_h) and those for consumers (subsidies s_h). In particular, such joint incentives increase both the probability that the seller mentions financial savings associated with the energy efficient water heater and the number of sales of the more energy efficient technology. By contrast, a consumer rebate without sales incentives does not lead the seller to disclose information about financial savings associated with the energy-efficient technologies, strategically marketing the energy-efficient version to a small minority of responsive consumers which ultimately results in undertreatment of high usage consumers.

In sum, subsidies for energy-efficiency interact with supply-side incentives, and given the lack of verifiability and liability on the market for energy-efficiency this may favor strategic pricing and information disclosure. In turn, further research should be directed at the design of subsidies in a market for credence goods and their impacts on prices. In Chapter II of this thesis, I study how subsidies for solar photovoltaic systems affect

---

[20] In addition to supply-side moral hazard, the authors also find evidence of demand side moral hazard, as insured patients tend to consult physicians more often.

[21] Pless and van Benthem (2019) discuss ways to prevent expert sellers from manipulating features of solar photovoltaic systems in order to appropriate subsidies. See also Aldy et al. (2018) and De Groote and Verboven (2019) for a discussion of how subsidies affect investment costs and electricity production.

their prices and their electricity output. The results show that larger subsidy levels are associated with increased costs.

3.4 Reputation and repeated interactions

In a context where trust matters, expectations about repeated interactions between expert sellers and consumers may encourage honest behavior. In line with this, Dulleck et al. (2011) find that providing consumers with information about interactions in previous rounds reduces overpricing, and increases the number of trades. Reputational concerns can potentially be relevant in the market for energy-efficiency, for example by providing trustworthy information about past behavior.[22]

Field evidence from car repairs by Rasch and Waibel (2017) suggests that garages in vicinity of a highway, and who are presumably less orientated towards repeated business, overprice more frequently. Similarly, Schneider (2012) finds that diagnosis fee for car repairs is significantly higher if consumer signals a one-shot interaction (stating to be moving away after the service and having moving boxes in the trunk). Note, however, that Schneider (2012) finds no evidence that signaling repeated business opportunities affect the quality of service, as undertreatment occurs with similar frequencies for consumers who signal single or multiple interactions.

In the context of energy-efficiency, investment decisions are infrequent, and at the individual level it is difficult to leverage reputational concerns to induce honest behavior. Nevertheless, the literature suggests that consumer learning and reputation building by expert sellers matters for investment decisions. In the market for cars, where consumers can learn about unobserved quality attributes through the experience of other consumers (e.g. through press reports), Heutel and Muehlegger (2015) show that adoption of high quality hybrid vehicles (the Toyota Prius) propagated a signal of high quality and increased the market share of all other models of hybrids. Conversely, the Honda Insight was perceived to be of lower quality, and its adoption had a negative impact on trust for that technology, leading to lower adoption rate for all other hybrid vehicles.

In the absence of verifiability, however, ex-post quality assessment for households is costly (e.g. by an independent auditor), which implies that bilateral feedback schemes are likely to be driven by subjective assessments and herding effects. Instead, what

---

[22] One example where bilateral feedback systems have successfully enforced trust are online markets for search goods (Tadelis, 2016).

is needed is a credible measure of whether a given seller delivered an economically viable technology (as per inequality 1). Such information can take the form of realized financial savings by previous consumers $p_e\hat{m}_i(\hat{e}_l - \hat{e}_h)$, for example. In line with this, Gillingham and Tsvetanov (2018) study a program in which households interested in performing an energy audit of their dwelling are provided with information on realized monetary and energy savings measured for other audited households. This provides credible information about the trustworthiness of the expert seller. In turn, households who have access to such information are more likely to carry out an audit themselves.

4 Discussion and conclusion

In this paper, we have investigated the credence component of energy-transforming technologies, arguing that the credence goods framework can be useful to further our understanding of the adoption of energy-efficient technologies and the associated energy-efficiency gap. Starting from a basic model of energy-efficiency investments, we identify how asymmetric information can lead to three types of supply side inefficiencies discussed in the credence goods literature, as well as low market participation. We highlight the difficulty to quantify energy savings, the associated failure of both verifiability and liability for many energy-transforming technologies, and the often observed discrepancy between ex-ante and ex-post energy savings. Taken together, this suggests that insights about inefficiencies inherent to credence goods, such as the baseline experiment of Dulleck et al. (2011), are useful for researchers and policymakers dealing with the market for energy-efficiency.

We then surveyed lab and field evidence on credence goods and relate key empirical findings to the literature on the adoption of energy efficient technologies. We can summarize the implications of our work for energy-efficiency policies along three main axes: (i) trustworthy certification schemes; (ii) subsidies in the presence of asymmetric information; and (iii) energy audits, reputation, and information on realized energy and financial savings.

First, we have highlighted that traditional energy labels and certification schemes involve a trust component. One implication is that third party verification of information and strict liability rules are necessary to ensure that labels provide decision-relevant information for consumers. In this context, we have discussed the role of competitor testing as an approach to detect fraudulent information provision. On the one hand, competing firms who comply with certification obligations have an incentive to make sure other firms comply as well. On the other hand, firms may be better informed than regulators as to how best perform relevant tests. Competitor testing therefore provides a promising approach to enforce liability in the domain of energy certification.

Second, evidence from credence goods markets suggests that subsidizing energy-efficient technologies is likely to increase markups, and therefore trigger overtreatment of consumers. In the context of energy efficiency, subsidies may thus lead consumers to invest in technologies with negative net present values. Consequently, supply-side adjustments to higher markups can be expected to affect the welfare of consumers in relation to subsidies for energy-efficient technologies. That being said, we note that empirical evidence on how prices for energy efficient technologies respond to subsidy policies remains thin, and further research in this area is warranted.

Third, separating diagnostic from treatment can reduce inefficient behavior by expert sellers, but the additional cost of audits can also reduce overall market efficiency. As a substitute, providing empirical evidence on realized energy savings and associated financial implications for already completed projects (as in Gillingham and Tsvetanov, 2018) may help consumers reach better investment decisions. Similarly, sharing information from ex-post audits through feedback platforms may help suppliers build a reputation. Promoting these relatively cheap sources of external information for consumers may help mitigate asymmetric information problems in the domain of energy-efficiency investments.

We close by emphasizing that research in energy economics may also benefit to the wider literature on credence goods. For example, the use of performance contracting has developed for specific investments and technologies, and a number of papers in the energy economics literature discuss the market for such contracts. But performance contracting can also be applied in other contexts, and Bester and Dahm (2017) discuss the possibility to spread payments to healthcare providers over time and condition these on patients' satisfaction with the treatment. Applying a credence goods framework to study specific energy-efficient technology may therefore be useful to broaden the scope of contributions made by energy economists.

Chapter II

Seductive subsidies? An analysis of second-degree moral hazard in the context of photovoltaic solar systems

This chapter is mainly based on a single-authored paper by Evert Reins. An earlier version of this paper has been published in the working paper series of UNINE IRENE (No. 21-03). It was presented by Evert Reins at the 12th Swiss Association of Energy Economics Student workshop, online (2020) and at the Annual Congress of the International Association for Energy Economics (online) 2021.

Abstract

This paper studies how subsidies for solar photovoltaic (PV) systems can lead to second-degree moral hazard - the impulse of installers to increase factors determining the total subsidies and/or transaction when consumers receive larger subsidy levels. Employing an instrumental variable strategy using plausibly exogenous variation in the size of subsidy levels to address concerns about self-selection of installers into specific subsidy levels, I quantify the impact of subsidy levels on the expected electricity output and transaction prices of PV systems in California. The results are consistent with hypothesized drivers of second-degree moral hazard as larger subsidy levels are associated with i) an increased measure of the expected electricity output leading to increased subsidies when third parties own the PV system and ii) increased transaction prices when consumers themselves own the system. The results further suggest that subsidy programs should verify the work of an installer, for example during mandatory field inspections, as these reduce second-degree moral hazard.

Keywords PV systems; Credence goods; Subsidies; Asymmetric information; Second-degree moral hazard.
JEL codes H23; H32; H76; D82; Q42; C26.

1 Introduction

Many countries use generous subsidy programs to accelerate the adoption of green technologies such as solar photovoltaic systems. To maximize the social and environmental value of each of the tens of billions USD spent on subsidy programs in the US and other countries (International Energy Agency, 2016), subsidy programs should be cost-effective. One key challenge to cost-effectiveness is that some characteristics of PV systems are subject to informational asymmetries typical for credence goods. In particular, installing PV systems consists of a complex arrangement of different technological components and working steps (Giraudet et al., 2018; Gillingham et al., 2016a) which many consumers deliberately leave to professional installers. Moreover, consumers face difficulties in verifying whether a system is installed and priced appropriately because the definition of a counter-factual relative to which the electricity output and price is measured is difficult (Giraudet, 2020; Lanz and Reins, 2021).

Therefore, professional installers may have incentives to exploit their informational advantage, leading to supply-side inefficiencies typical for credence goods, including overcharging for services or technological components (Giraudet, 2020; Lanz and Reins, 2021). The literature on credence goods further suggests that third-party reimbursements may cause second-degree moral hazard and thereby increase supply-side inefficiencies (Balafoutas et al., 2017). Because subsidies reduce the transaction price paid by consumers, installers may be *more* inclined to increase the transaction price of PV systems (Kerschbamer et al., 2016; Huck et al., 2016; Balafoutas et al., 2017, 2020). [1] In addition, when the total amount of subsidies received is determined by self-reported values on system characteristics, installers may have incentives to exaggerate such values in order to increase the overall amount of subsidies received.

In this paper I study how subsidies may trigger second-degree moral hazard and hence i) increase the total amount of subsidies received and ii) increase transaction prices of PV systems. For this purpose, I use data from the California Solar Initiative (CSI) which

---

[1] Giraudet et al. (2018) discuss how demand- and supply-side moral hazard affect the installation quality of home energy retrofits. They argue that under an energy savings contract, where the installer reimburses the consumer for shortcomings of energy savings and thus reduces his marginal costs of energy consumption, the consumer may increase his energy consumption which can be interpreted as demand-side moral hazard. Furthermore, Giraudet et al. (2018) refer to supply-side moral hazard in situations where installers exploit the credence component of energy retrofits and provide installations with poor quality. Although subsidies reduce the total cost (instead of the marginal cost) of adopting energy efficient technologies, second-degree moral hazard, as studied in this paper, is related and describes how subsides may further increase supply-side moral hazard because consumers do not bear the full total costs of their investment.

is the largest solar subsidy program in California. Idiosyncratic characteristics of the CSI and related data on subsidized PV systems make the program particularly relevant for this research. First, the CSI offers regional and chronological variation of subsidies enabling the identification as to how the expected electricity output as measured by the design factor of a system and transaction prices depend on the magnitude of received subsidies. Specifically, the CSI provides subsidies to consumers in three different energy supply companies (or investor- owned utilities, IOU) following the aim to generate a total of 1940 megawatts (MW) capacity installed in new PV systems. The subsidy level available to consumers is categorized in ten predetermined steps where the transition from one subsidy level step to the next is determined by the of cumulative capacity of MW installed within an IOU.

Second, to calculate the total amount of subsidies a consumer receives, the CSI program uses self-reported data on expected electricity output (or hereafter the design factor) of a system. In particular, consumers have to report specific system characteristics, such as its location, shading, orientation as well as the make and model of installed PV modules and inverters. However, only a fraction of PV systems is subject to a mandatory field inspection where the accuracy of these characteristics is verified. Increased verification is a well studied countermeasure to supply-side inefficiencies in markets for credence goods and I can identify systems where the applicant knew that the system characteristics were verified during a field inspection by the CSI, allowing me to assess how such verification affects second-degree moral hazard.

In the empirical analysis, I quantify the impact of subsidy levels on the design factor and transaction prices, using variations of subsidy levels afforded by the design of the CSI program. Following Pless and van Benthem (2019), I estimate linear models using a rich set of fixed effects (FE). Specifically, I employ fixed effects along four axes: i) installer FE to capture time-invariant installer specific characteristics such as their market power, ii) month of installation FE to capture national demand shocks and general time trends for hardware prices, iii) regional FE to capture local differences in demand and competition among installers and iv) make and model of modules and inverters to control for unobserved differences in the installed technology, such as their quality.

I further make use of an instrumental variable strategy to address potential concerns about the endogeneity of actually implemented subsidy levels. The actually received subsidy levels differ from predetermined subsidy levels for some systems and I cannot rule out that installers were able to influence actually received subsidy levels, thereby self-selecting into specific subsidy levels. In particular, some PV systems receive weighted averages of up to 4 different subsidy level steps in contrast to sharp and monotonic decreases of subsidy levels as determined by the CSI design.

To address this concern, I exploit plausibly exogenous variation of predetermined subsidy levels to instrument actually received subsidy levels. In this context, the validity of this instrument rests on two assumptions. First, the predetermined subsidy levels need to be correlated to the actually implemented subsidy levels. This assumption is likely to hold because the predetermined subsidy steps are the predominant factor determining received subsidy levels and the differences between predetermined and actually received subsidy levels is small.

Second, the exclusion restriction implies that the ex ante-determined subsidy level steps do not affect the design factor and transaction prices other than through the actually received subsidy levels. Again, this assumption is likely to hold as even large installers could not influence the total capacity installed within an IOU. I further include the above set of fixed effects in the first stage and thereby control for any between installer, between month, between IOU and county and between technology factors that potentially link subsidy steps to the design factor and/or transaction prices of systems.

I study heterogeneity of supply-side inefficiencies along three dimensions that have been highlighted as important factors driving second-degree moral hazard (see Balafoutas et al., 2017, for a related discussion). As a first dimension, I study whether second-degree moral hazard is increased if a system is third-party owned (TPO) and hence the installers directly receive the subsidy as opposed to home-owned (HO) systems where the consumer receives the subsidy. As a second dimension, I study increased verification of installations and their potential to prevent second-degree moral hazard (Dulleck and Kerschbamer, 2006; Dulleck et al., 2011). To this end, I exploit a CSI rule imposing a mandatory field inspection for the first two PV systems installed by each installer. Finally, I study how second-degree moral hazard depends on whether a system is owned by a commercial, residential, non-profit or governmental consumer.

The empirical analysis shows evidence suggesting that second-degree moral hazard is highly relevant in the context of upfront subsidies. First, I find that a one dollar increase of upfront subsidies is associated with a statistically significant 0.5 percentage point increase of the design factor of residential TPO systems. Such an increase is for example equivalent to reporting a five degree difference in the module tilt toward the optimal tilt. When systems are subject to a mandatory field inspection, there is no statistically significant association, suggesting that increased verification prevents installers from reporting exaggerated system characteristics. I do not find any significant marginal effect of the subsidy level on the design factor of HO systems.

Concerning transaction prices, I do find a significant marginal effect of the subsidy level of HO systems for all consumer sectors, suggesting that a one dollar increase of the subsidy level increases the transaction prices by 3.5 to 7.5 percent at the mean subsidy level. For TPO systems, I only find such an effect for governmental consumers. This effect is however very large as a one dollar increase of the subsidy level is associated with a 25 to 50 percent increase of the transaction price per Watt at the sample mean. Overall, these results are in line with the intention of installers to increase their short-term cash flow, because TPO installers receive the total subsidies and HO installers receive the transaction price directly after the installation.

These findings contribute to three different kinds of research avenues. First, Davidson and Steinberg (2013) and Podolefsky (2013) have documented that some TPO installers inflate the transaction price of residential PV systems to reap larger tax credits. Yet, Pless and van Benthem (2019) somewhat surprisingly find that pass through of residential TPO systems receiving upfront subsidies is higher and attribute this effect to imperfect competition on the market for TPO systems in combination with a sufficiently convex demand curve. The results in this paper suggest that TPO installers do increase the total amount of subsidies received and at the same time, do not adjust the transaction prices. Therefore, second-degree moral hazard provides a parallel explanation for the over-shifting of subsidies found in Pless and van Benthem (2019) as TPO installers do not increase transaction prices after increasing the total amount of subsidies received. Potentially, competitive pressure in TPO markets may drive installers to increase cash-flow by inflating short-term payments via larger subsidy transfers and then acquire more new consumers with lower transaction prices.

Second, results from the literature assessing the optimal subsidy design to increase adoption and cost-effectiveness of subsidy programs show that consumers significantly discount future subsidy payments and that upfront subsidies are a cheaper way to foster the adoption of PV systems than feed-in subsidies (Burr, 2016; Feger et al., 2017; De Groote and Verboven, 2019). These studies do however not account for second-degree moral hazard associated with upfront subsidies and my analysis suggests that the final amount of upfront subsidies received should not be based on self-reported data on the expected electricity output of a system, unless this data is verified.

Finally, this paper substantiates that stylized findings from other credence goods markets are relevant for the market of energy-transforming technologies (Giraudet, 2020; Lanz and Reins, 2021). My analysis is the first to document evidence of second-degree moral hazard in the context of energy-transforming technologies. In line with earlier discussions on the potential of strict verification to reduce opportunistic behavior (Dulleck and Kerschbamer, 2006; Dulleck et al., 2011; Balafoutas et al., 2013), I find that mandatory field inspections can limit second-degree moral hazard related to increased total subsidies and also transaction prices. The results in this paper further confirm heterogeneity of supply-side inefficiencies depending on who bears their costs (Balafoutas et al., 2013; Gottschalk et al., 2020).

This paper proceeds as follows: in Section 2 I discuss the credence nature of energy-transforming technologies, second-degree moral hazard, and resulting consequences for the design factor as well as transaction prices of PV systems. Section 3 describes the CSI program. In Section 4, I summarize the data and explain the identification strategy. I present associated results in Section 4 and conclude in Section 5.

2 PV systems, supply-side inefficiencies and second-degree moral hazard

I illustrate the implications of the credence nature of PV systems in the context of the CSI program building on the framework of Dulleck and Kerschbamer (2006). I assume that there are two types of PV systems: those with high quality technological components q_h and those with lower quality technological components q_l. The electricity output of system i, $V_i(q_i, l_i, d_i)$ is increasing in the quality of technology, the quality of labor exerted during the installation l_i and the systems design factor d_i. The design factor is a measure of the system's real world potential for electricity output accounting for the system's technological components, its mounting method, orientation, tilt, azimuth, and shading as well as the solar irradiation at its location.

The installer faces a cost for installing a system which increases in both, the cost of technology and the cost for labor (i.e. $c_i(q_i, l_i)$). Then consumers pay the transaction price for the system which is increasing in labor and hardware costs $p_i(c_i)$. The installer's benefit from putting up a PV system hence equals the difference between the transaction price and costs of provided hardware and labor $\pi_{installer} = p_i(c_i) - c_i(q_i, l_i)$. The consumer's benefits from investing in a PV system can be expressed as $\pi_{consumer} = V_i(q_i, l_i, d_i) - p_i(c_i) + s_i$. In this equation, p_i is the (pre-incentive) transaction price determining the amount of money the installer receives for setting up the system and $p_i - s_i$ is the post incentive price determining what the consumer actually pays to set up his system accounting for the received subsidies.

The credence nature of PV systems implies that there is asymmetric information on q_i, l_i and d_i (see Giraudet et al., 2018, for a similar assumption). In particular, there is a significant cost to verify whether the self-reported information on the design factor is correct, which technology was installed and whether the technology was mounted and wired in an appropriate way. This implies that an installer has some margin with respect to determining the design factor and also transaction prices. If an installer knows that his work will not be verified and if he only cares about his own profits, he has the incentive to increase transaction prices. Such behavior can be expressed with $\sigma \in [0, 1]$ where $\sigma = 1$ indicates exploiting asymmetric information to a maximum leading to maximized total subsidies and transaction prices while $\sigma = 0$ is equivalent to the fair transaction price if there was no asymmetric information.

Next, evidence in markets for credence goods suggests that increased verification measures imposed to detect and punish supply-side inefficiencies may change the installer's behavior (Dulleck and Kerschbamer, 2006; Dulleck et al., 2011; Balafoutas et al., 2013). Let γ denote the probability of detecting supply-side inefficiencies σ and t denote related punishment. Intuitively, larger verification of the installer's work may work as a threat to lose financial and/or reputation status. The larger γ and t, the larger the expected disutility from supply-side inefficiencies. The CSI program administrators demand that the first two systems installed by each installer are subject to mandatory field inspections where the system's characteristics and functionality are verified. Such mandatory field inspections increase verifiability of the installers' work and allow me to estimate their effect on the design factor and transaction prices.

Moreover, it has been shown that agents in markets for credence goods care for the consumer's benefits, suggesting that installers have some form of social preferences represented by λ (see for example Kerschbamer et al., 2017; Kandul et al., 2020). Looking at the active market for PV systems, it seems plausible that a large heterogeneity in λ exists, implying that many installers care for the consumer's benefits and therefore provide flawless services (see for example Kerschbamer et al., 2017). At the same time, differences in the consumer owning the system may affect λ. When installers think about who bears the consequences of supply-side inefficiencies, they may for example want to reduce the burden for residential consumers whom they personally know (i.e when λ is large) compared to more abstract entities with several stakeholders and financiers such as governmental and commercial consumers (see Balafoutas et al., 2017, for a discussion of how social distance may affect the behavior of installers).

Adding these insights to the framework of Dulleck and Kerschbamer (2006), the objective of the installer can be written as follows:

$$\pi_i = p(q_i, l_i) - c(q_i, l_i) - \gamma t\sigma + \lambda(V_i(q_i, l_i) - p(q_i, l_i) + s_i).$$

The literature on credence goods further suggests that an installer may alter his behavior conditional on the magnitude of the subsidy the consumer receives. In particular, second-degree moral hazard in the context of PV systems implies that installers *further* increase the transaction prices of PV systems when consumers receive larger subsidies.

Importantly, the CSI program is designed such that the total amount of upfront subsidies is increasing in the design factor of systems $s_i(d_i)$ which is multiplied with the subsidy level and the system size to determine the total amount of the upfront subsidy. Second-degree moral hazard may therefore imply that installers also increase the design factor of systems to maximize the total amount of subsidies received.

3 The California Solar Initiative program

The California Solar Initiative (CSI) program provides a useful setup where predetermined subsidy levels are assigned to different consumers, enabling me to analyze the associated relation with the design factor and transaction prices of PV systems. This section starts with a general description of the program. Afterwards, I describe which features of the CSI are exploited to study second-degree moral hazard related to the subsidy recipient, increased measures of verification and differences in system ownership. Finally, I describe the main outcome variables of the empirical section: the design factor and the transaction price.

3.1 Program description

The CSI subsidy program was rolled out in 2007 using a budget of $2,167 million for the goal to install 1940 MW within 10 years. All consumer sectors could apply for the program including residential, commercial, government and non-profit consumers. The subsidy studied in this paper ("*Expected Performance Based Buyout*", *or EPBB*) is intended for residential and small business consumers installing a system with less than 30 kW and takes the form of a one time lumpsum payment. The size of the lumpsum payment is determined by multiplying the subsidy level with the design factor and system size.

The subsidy level available to consumers is determined by the cumulative capacity of already installed systems within the IOU of the consumer. Once a certain threshold of cumulative MW in an IOU is passed, the subsidy level decreases. In particular, the

Table 1: CSI subsidy levels

MW Step	**MW in step**	**Residential/ Commercial**	**Gov't/ Nonprofit**
1	50	n/a	n/a
2	70	2.5	3.25
3	100	2.2	2.95
4	130	1.9	2.65
5	160	1.55	2.3
6	190	1.1	1.85
7	215	0.65	1.4
8	250	0.35	1.1
9	285	0.25	0.9
10	350	0.2	0.7

Notes: Subsidy levels in $ per Watt. Extract from Table 4 of California Public Utilities Commision (2017).

CSI provides subsidies to consumers in investor-owned utility territories of Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE), and San Diego Gas & Electric (SDG&E). Table 1 provides an overview of the subsidy levels as per the design of the CSI. After the first 50 MW in each IOU have been attributed under another program (Lilly and Simons, 2006), passing the predetermined threshold of MW installed leads to a sharp and monotonic decline of subsidy levels in the IOU.

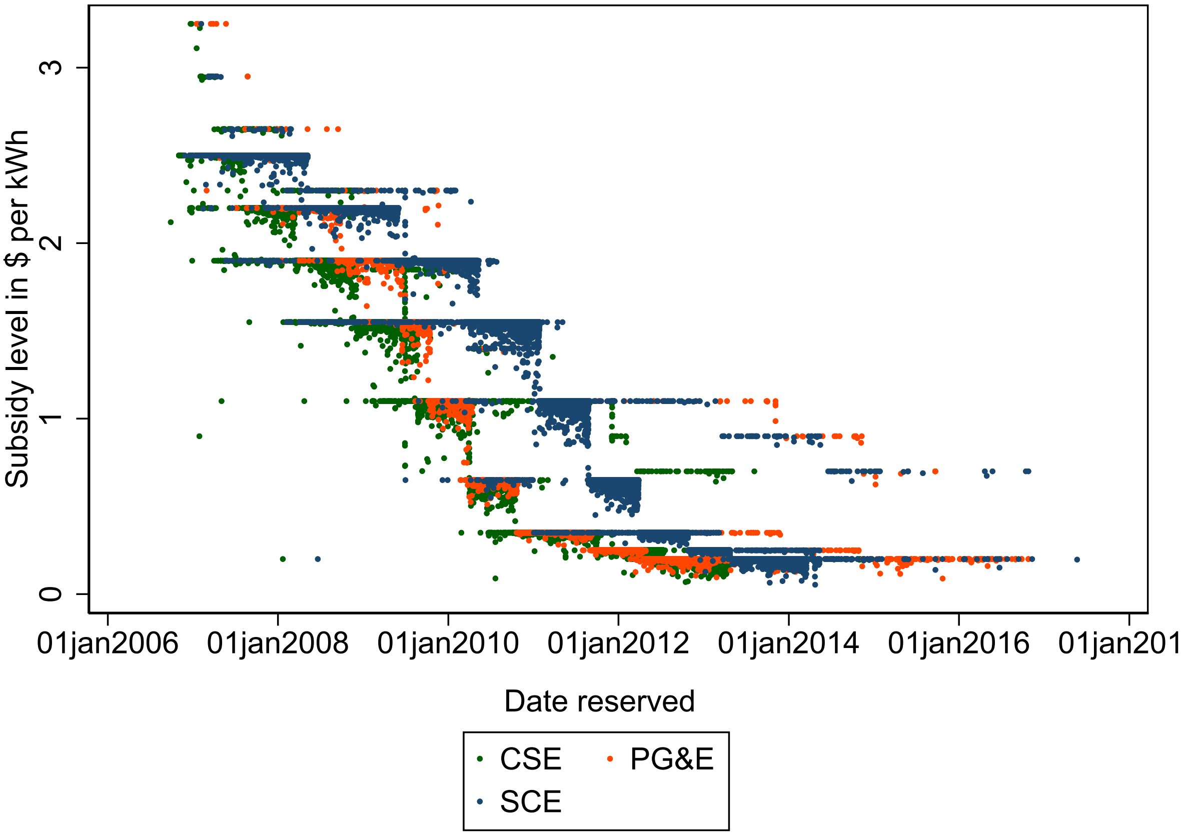
The actual implementation of subsidy levels however differed from the theoretical design. Figure 1 provides an overview of the implemented subsidy levels across IOUs and time. One can see that for example in January 2010, many different upfront subsidy levels were attributed to different PV systems in all IOUs. Contrasting the unique subsidy levels in each MW step, some systems receive weighted averages of up to four different predetermined subsidy levels. In addition, subsidy levels do not monotonically decline in time, but some systems which have applied in the same IOU at a later point in time receive a higher subsidy rate.[2]

While the subsidy levels attributed to systems are predetermined by the MW step, the total subsidy amount is increasing in the system size. Accordingly, installers can further increase the total subsidy amount a consumer receives by installing larger systems. To avoid the installation of unreasonably sized systems the CSI imposed rather strict limitations to determine a system's size and to ensure that a system is sized such that it optimally serves the consumer's needs. First, a system should primarily offset the applicants own energy consumption, meaning that the annual expected electricity

---

[2] The CSI handbook does not provide an explanation for these observations which contrast the theoretical design of the CSI. Presumably, the fact that some systems receive a weighted average of several subsidy levels could be either due to cancellation of systems and liberated capacity under a subsidy step which was already exhausted or an adjustment of CSI subsidies if systems receive other benefits (Hughes and Podolefsky, 2015).

Figure 1: Evolution of subsidy levels



Notes: Subsidy levels over time and IOU's.

output must not be larger than the sum of energy consumption within the last twelve months. Second, no applicant may receive a total amount of subsidies that exceeds the transaction price of the system. Third, there is a cost cap for applications implying that the transaction price per Watt may not be larger than the 12 month rolling average of the transaction price per Watt of other systems plus one dollar.

Studying the distribution of systems around two arbitrary thresholds provides information on whether installers strategically influence the size of systems. First, systems smaller than five kW were not required to submit a substantiation of the system size when applying for the CSI. Second, systems smaller than ten kW did not have to pay an application fee.[3] Figure A1 in Appendix A presents the size distribution of PV systems in the range between zero and 30 kW in the upper panel, between four and six kW in the lower left panel and between nine and eleven kW in the lower right panel. There are not disproportionately many systems sized just below five or ten kW, affirming that the system size is determined by the consumer's needs rather than strategic considerations (see also Gillingham et al., 2016a; Pless and van Benthem, 2019, for similar conclusions on the sizing of PV systems).

3.2 Subsidy recipients

Instead of buying a PV system, CSI consumers can choose to lease a PV system from a third party (Podolefsky, 2013; Pless and van Benthem, 2019).[4] In this case, TPO installers pay the installation costs and receive the final subsidy (i.e. they directly receive s_i Equation 2).

The US treasury department has investigated the pricing of some TPO installers, as these were accused of increasing fair-market values in order to reap larger tax credits (Trabish, 2013). Because TPO installers finance the upfront installation costs and the transaction price to installers is paid in form of dispersed leasing rates, TPO installers

---

[3] For other system sizes the application fee equals 1250 USD for systems up to 30kW. Note that this fee is refunded once the system is installed.

[4] If consumers choose to lease a system, they can decide between a pure leasing contract or a power purchase agreement (PPA). In a pure leasing contract, the consumer pays a monthly leasing rate to the third party and owns the electricity output. In a PPA contract, the consumer pays a monthly rate for his electricity consumption and the third party owns the electricity output. The contract types mostly differ with respect to who is entitled to the benefit of excess output fed into the system. Under either contract type, the third party pays for the installation and maintenance of the system and consumers hence do not bear the upfront costs (see Davidson et al., 2015, for a detailed discussion of pure lease and PPA contracts). In the dataset, I can identify the systems owned by a third party but I cannot identify whether they have a leasing or PPA contract.

may have incentives to increase short-term cash flows in order to increase their market value (Salzman, 2013). As the CSI subsidies are paid directly after the installation of a system, TPO installers can increase their short-term cash flow by increasing the total amount of subsidies received. As discussed by Pless and van Benthem (2019), such a business strategy may affect estimates of subsidy pass-through, because inflated total subsidies would artificially decrease the post incentive price and therefore falsely be attributed to a larger pass-through.

3.3 Increased verification

Following the CSI rules, the first two PV systems installed by each installer are subject to an onsite field inspection which serves the goal to detect differences between the onsite technical calibrations of the system and those stated in the application form to calculate the design factor.[5] In particular, mandatory field inspections thus include checking that equipment is installed as documented in the application (i.e. quantity and make of modules and inverters, a systems tilt, azimuth, shading and standoff height) as well as whether the system is operational and its electricity output is reasonable. Finally, if subsidy payments resulting from onsite inspections and those calculated in the application form documentation differ by more than 10 percent, the PV system and its installer can be dismissed from the program.

This rule is public knowledge and thus known by installers. Mandatory field inspection increase the probability γ of detecting supply-side inefficiencies and installers face commercial consequences after detection (i.e. $t > 0$). In turn installers may limit exaggerating the design factor and/or increasing transaction prices in order to prevent financial and reputational consequences in case of detection (cp. Balafoutas et al., 2013; Giraudet et al., 2018, who find that increased verification reduces increased prices and supply-side inefficiencies are specifically pronounced in domains defined as *hard to observe*).

3.4 Sector of consumer

Furthermore, installers in the sample face commercial, residential, non-profit and governmental consumers. This enables me to study differences behind the entities owning PV systems which differ with respect to financial resources and social distance

---

[5] The CSI further has the right to audit additional systems according to its own assessment. These audits are either performed online, via telephone or onsite.

(i.e. heterogeneity in λ). Following Balafoutas et al. (2017), installers may be more inclined to increase transaction prices when consumers are perceived as wealthier and the financial consequences are borne by an anonymous entity compared to a residential consumer with whom interaction is more direct and personal.

Evidence on distributional preferences of agents in markets for credence goods suggests that supply-side inefficiencies are reduced when they have larger financial consequences for the consumer (Kandul et al., 2020). If installers perceive non-profit and residential consumers as less financially endowed and therefore have a higher valuation for their benefits (i.e. a larger λ) compared to commercial and government consumers, one would expect to observe differences in second-degree moral hazard depending on the consumer sector.

3.5 Measures of the design factor and transaction price

To document second-degree moral hazard in the context of PV systems this paper analyses the design factor and transaction prices of PV systems under different subsidy levels. The design factor is calculated by the CSI, based on the following criteria reported by the PV system's applicant: the zip code and IOU of the installation location, the sector of the applicant, the make, model and number of PV modules and inverters, the mounting method, the tilt and azimuth of the PV system and the shading of the system including a precise measure if there is shading. The CSI then calculates the expected production of the proposed system based on the reported characteristics and compares it to a reference system. In particular, this comparison includes i) a design correction to account for differences in tilt and azimuth, ii) a geographic correction to account for differences in the location with respect to temperature and solar irradiation at the respective zip code, and iii) an installation correction to account for differences in the mounting method relative to laboratory test conditions. While the zip code as well as make, model and number of technological components may be easier to verify and applicants would therefore need to deliberately misreport in order to increase the design factor, applicants may exploit the asymmetric information and report exaggerated measures of features which are hard to observe, such as a system's shading, tilt and azimuth.

The transaction price of PV systems is the second outcome variable of interest. In line with technical conventions it is divided by the system size. In particular, the system's nameplate rating is used to determine the transaction price per Watt as this measure reflects the system's electricity generating potential under test conditions (see

Podolefsky, 2013; Hughes and Podolefsky, 2015; Dong et al., 2018; Pless and van Benthem, 2019, for a similar procedure).[6] The CSI data provides the transaction price for each system, which includes costs for the technological components, construction and installation costs, engineering and design costs, interconnection cost as well as warranty and maintenance costs.[7] While the costs for technological components may be easily verifiable, the idiosyncratic environment of each system demands specific installation and maintenance work-steps where installers likely have some range to exploit with regard to pricing.

Importantly, earlier literature has assessed the transaction price of HO systems as a reliable measure of what a consumer pays to the installer before subsidies. For TPO systems however, this measure has been deemed inconsistent because the actual transaction price depends on the explicit contract details such as monthly payments, term lengths and upfront payments (Pless and van Benthem, 2019). I therefore conduct the analysis of transaction prices per Watt of HO and TPO systems separately (see Section 5.2). I decide to keep the transaction prices of TPO systems as a benchmark for comparisons and make a disclaimer that all qualitative and quantitative conclusions related to the transaction prices of TPO systems are indicative.

4 Data and empirical strategy

I first provide a summary of the data in Section 4.1 and then present the identification strategy to investigate supply-side inefficiencies and second-degree moral hazard in Section 4.2.

4.1 Data summary

The information on the applicants of each PV system provides a rich set of system characteristics. Table 2 presents summary statistics of the data for each year of the

---

[6] The CSI data report three different measures of a system's size. The nameplate measures the electricity generating potential under standard test conditions. The CEC-AC rating accounts for differences at the respective location of the system such as wind speed and ambient temperature. The CSI rating equals the CEC-AC rating multiplied by the design factor of the system.

[7] All systems in the sample are eligible to receive investment tax credits (ITC). These take the form of a 30 percent tax credit on the transaction price which was granted to all PV systems installed in the US from 2006-2019. Adjusting the transaction price for the ITC is akin to a linear transformation of the variable, as all systems installed during the sample period receive the same tax credit. Consequently, this procedure would not affect the results. See Pless and van Benthem (2019) for further information.

sample time (2007 to 2016).[8]

Table 2: Summary statistics

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Mean subsidy level ($/W)	2.40	2.02	1.58	1.02	0.59	0.30	0.21	0.20	0.21	0.21
Min subsidy level ($/W)	0.90	0.20	0.65	0.09	0.21	0.07	0.06	0.05	0.09	0.15
Max subsidy level ($/W)	3.25	2.65	2.30	2.30	1.55	1.10	1.10	0.90	0.70	0.70
Mean price per Watt ($/W) (HO)	8.2	8.1	7.8	7.1	6.8	5.5	4.7	4.6	4.2	4.0
Mean price per Watt ($/W) (TPO)	8.2	9.8	7.7	7.1	6.4	5.4	5.0	4.5	4.7	4.9
Mean size in kW	6.4	5.7	6.1	5.8	5.6	6.1	6.5	6.8	7.8	8.1
Mean number of modules	34	30	31	27	25	24	25	26	29	29
Mean number of inverters	1	2	3	4	5	6	8	9	11	10
Mean design factor (in pct.)	94.48	94.04	94.46	94.28	94.67	94.29	94.48	95.04	94.73	95.99
Commercial (1,943)	2.8	2.8	1.8	1.8	1.0	0.8	0.8	2.3	4.4	10.6
Government (374)	0.6	0.5	1.1	0.4	0.8	0.0	0.0	0.1	.	0.6
Non-profit (565)	1.1	0.7	0.4	0.5	0.3	0.2	0.2	0.9	3.2	2.5
Residential (135,768)	95.5	96.0	96.6	97.3	98.6	98.9	98.9	96.7	92.3	86.2
Observations (138,650)	6477	9701	13344	18994	21692	31691	30416	5677	498	160
Observations (HO, 70,914)	6019	8308	11429	13120	10179	8918	10134	2407	297	103
Observations (TPO, 67,736)	458	1393	1915	5874	11513	22773	20282	3270	201	57

Notes: Summary statistics for 2006 and 2017 are not reported because there were only 81 applications in these years.

The first three rows show the mean, minimum and maximum subsidy level in a given year. In line with Figure 1, subsidy levels vary considerably within each year. The next two rows show the average transaction price per Watt, for HO and TPO systems separately because of potentially inconsistent values for TPO systems (see Section 3.5 for further details). Importantly, the transaction price per Watt is declining over time for both HO and TPO systems. This trend is in line with a decrease of hardware costs in recent years. Because the CSI subsidy levels also decrease over time, controlling for changes in time-varying factors affecting the transaction price and design factor of PV systems is crucial when estimating second-degree moral hazard.

Next, the size in kW and the number of modules installed show no specific time trend. The number of modules installed per Watt is decreasing which is in line with the trend that the peak kW (i.e. the maximum kWh generated per module) has increased over time and fewer modules are necessary to reach a given level of electricity output. Also, the total number of inverters installed is increasing over time. Solar inverters are the primary cost drivers of a PV system and their hardware costs have significantly decreased over the sample period. In combination with the observation that the ratio of inverters per module installed also increased, this suggests that early adopters of PV technology limited the installation of installers to cap hardware costs.

---

[8] Note that I drop PV systems which have not been installed at the time of the data access. Also, systems without entries for the subsidy level, transaction price or date of reservation were dropped in the data.

The design factor stays constant over the sample period, indicating that neither *low-hanging fruits* with a particular large design factor, nor systems with a particularly bad potential to generate electricity output entered the CSI program early. The next four rows show the distribution of the consumer sector. The systems installed in our sample are predominantly owned by residential consumers, followed by commercial, non-profit and government consumers which is in line with the intended allocation of EPBB to smaller and residential consumers. Finally, one can further observe a strong growth of TPO systems during the sample period.

4.2 Identification strategy

To estimate the association of subsidy levels and the design factor and transaction prices per Watt, I employ regression specifications adapted from Pless and van Benthem (2019). When the outcome variable is the design factor, the regression specification can be written as follows:

$$Y_i = \alpha + \beta_i s_i + X_i \phi + \varphi_u + \varsigma_f + \chi_s + \omega_c + \mu_t + \epsilon_i \tag{1}$$

where Y_i denotes the design factor of system i, and s_i denotes the subsidy level of system i. In addition, I account for entered system characteristics used to calculate the design factor by controlling for the number of modules and inverters $X_i \phi$ and further employing IOU fixed effects φ_u, technology fixed effects ς_f indicating the make and model f of modules and inverters installed in system i to control for quality differences, and sector fixed effects χ_s to control for differences in subsidy levels in the respective sector (see Table 1).

I also make use of installer fixed effects ω_c to eliminate potential bias at the installer level such as measurement errors of the system characteristics. Further, I employ μ_t which is a dummy variable for the month t in which system i was installed to control for the development of the design factor over time. This is important because one could argue that consumers with a particularly poor environment for solar electricity generation have opted in the CSI program early, because only high subsidy levels make the investment for such consumers profitable (see Globus-Harris, 2020; Gilbert et al., 2019, for a related discussion of additionality effects). The inclusion of monthly fixed effects prevents me from misinterpreting the association of subsidy levels and the design factor as second-degree moral hazard when it can be actually attributed of additionality.

Finally, ϵ_i denotes a random error term and standard errors are clustered at the zip code level to correct for potential correlation of data errors within regional CSI offices (Podolefsky, 2013; Pless and van Benthem, 2019).

When the outcome variable Y_i is the transaction price per Watt, I include additional variables in the vector $X_i\phi$ to control for competition and other aspects of the market structure affecting the pricing and supply for PV systems. First, Pless and van Benthem (2019) find evidence suggesting that the market for solar photovoltaic systems in California is affected by imperfect competition, in particular heterogeneous degrees of market power across firms. To control for the development of market power of a firm over time, I follow Dong et al. (2018) and compute the rank of the respective installer in terms of total installations within the zip code at month t. To further account for local industry concentration and competition, I follow Gillingham et al. (2016a) and compute the year-zip code level Herfindahl-Hirschmann Index (HHI) using the number of installations of a respective installer in the zipcode to calculate his market share in the respective year. Second, Bollinger and Gillingham (2019) show that the installation costs of solar photovoltaic systems in California decrease in the experience of installers, which is attributed to learning-by-doing. To control for the experience of installers, I therefore include a measure of many PV systems each installer has installed previous to system i.

I furthermore include the same FE as described above when the outcome variable is transaction price per Watt. In particular, using installer FE in this context is important to address unobserved differences at the installer level, such as increased market power due to more successful marketing campaigns or increased bargaining power when negotiating transaction prices with the consumer and/or lower prices for technology input with the manufacturer.

There is, however, a potential issue with specification 1 because the actual received subsidy levels differed from the predetermined subsidy levels for some observations for reasons which were not explained in the CSI program (California Public Utilities Commision, 2017). I can thus not rule out that installers are able to influence the subsidy level and therefore self-select into specific subsidy levels. To address this concern, I exploit plausibly exogenous variation of the predetermined subsidy level as part of an instrumental variable strategy. In the first stage, I instrument the actually received subsidy level with the predetermined subsidy level depending on the cumulative MW installed within an IOU (see Table 1). Because the actual allocation of subsidy levels was mostly in line with the predetermined schedule, predetermined subsidy levels are a good predictor of actually received subsidy levels.

Further, the exclusion restriction requires that the instrument affects the design factor and transaction price per Watt of systems only through the subsidy level. Importantly, the exclusion restriction is conditional on a set of control variables and I include the above mentioned fixed effects in the first stage, thereby controlling for any between installer, between month, between IOU and county and between technology factors that potentially link subsidy steps to the design factor and transaction price of systems. In particular, the predetermined subsidy level decreases with the cumulative MW installed within an IOU and the instrument is therefore to some extent related to the time course of the sample period. As shown in the following, including monthly fixed effects is thus of high importance to validate the exclusion restriction.

The design factor is essentially a measure of a system's expected electricity output affected by solar irradiation, tilt, azimuth and mounting method at the consumer's site relative to optimal laboratory conditions (see section 3.5 for further details). Self-selection of consumers with a low design-factor applying in early stages of the CSI in order to secure large subsidy levels could possibly create a direct link of the instrument to the design factor and thus invalidate the exclusion restriction. However, the inclusion of monthly fixed effects in the first and second stage isolates any potential co-movement of the cumulative capacity of MW installed and the design factor.

One could further imagine that the cumulative capacity of MW installed within an IOU depends on the transaction price, thus causing a reverse causality between the outcome and the instrument. Again I use the within monthly predetermined subsidy level as an exogenous shifter of the actually received subsidy level. It is highly unlikely that any within monthly-changes of the cumulative capacity installed depend on changes in the transaction prices, as these decline slowly and gradually over the sample period (see Pless and van Benthem, 2019, for a similar arguement).

Formally, the received subsidy level (see Figure 1) is instrumented with the predetermined subsidy level depending on the cumulative MW installed as presented in Table 1:

$$Z_i = predetermined\ s_i \tag{2}$$

Consequently, the first stage regression can be written as:

$$s_i = \eta + \theta Z_i + X_i \tau + \vartheta_u + \varrho_f + \iota_s + \xi_c + \kappa_t + \nu_i. \tag{3}$$

Using this instrumental variable approach, the second stage estimate β accounts for potential endogeneity of the actually received subsidy levels. The estimate is further based on within month, within IOU, within county, within installer and within technology variation of subsidy levels. Controlling for additional factors which potentially influence the design factor and transaction prices per Watt, I interpret β as the causal relation between a one dollar subsidy increase and associated changes of the design factor and transaction prices per Watt of PV systems.

I then study whether the association between subsidy levels and design factor/ transaction price per Watt is affected by increased verification of the installer's work and the ownership of systems. For this purpose, I interact the subsidy level s_i in specification 1 with a variable indicating whether i) a system is subject to a mandatory field inspection, and ii) the system is owned by a commercial, government, non-profit or residential consumer. This procedure requires that I instrument each interaction term with the predetermined subsidy level interacted with the category of the indicator variables, resulting in several first stage regressions. To ease the interpretation of the interaction terms, I further center the subsidy level variable around its mean. Hence, interaction terms can be interpreted as the association of subsidy levels and the design factor/ transaction prices per Watt at the mean subsidy level of the sample.

I complement this identification with a set of robustness checks. Hughes and Podolefsky (2015) as well as Pless and van Benthem (2019) note that consumers could to some extent anticipate subsidy step transition dates and therefore speed up the application process to receive higher subsidy levels. I therefore follow Hughes and Podolefsky (2015) and Pless and van Benthem (2019) and drop systems which applied in the vicinity of two weeks before and after a subsidy level drop. I then apply the instrumental variable strategy this subset of data.[9]

---

[9] The possibility that consumers are able to decide on which side of the threshold for a subsidy step in combination with the irregularities concerning the actually received subsidies impedes me from using a regression discontinuity design. The instrumental variable strategy however mimics the first stage regressions which one would have performed to determine abrupt subsidy level changes in the vicinity of threshold for a transition to a next subsidy step.

Table 3: Design factor of TPO systems

	All obs. included					
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Subsidy level	0.359*	0.471**	0.360*	0.471**	0.707	1.315
	(0.144)	(0.164)	(0.144)	(0.164)	(0.776)	(0.714)
Field inspection (FI)						
FI = 1 × Subsidy level			−0.602	−0.421		
			(0.445)	(0.467)		
FI = 1			0.708	0.690		
			(0.423)	(0.422)		
Sector						
Government × Subsidy level					−1.086	−2.034
					(2.228)	(3.103)
Non-Profit × Subsidy level					0.795	1.324
					(0.939)	(1.388)
Residential × Subsidy level					0.360*	0.470**
					(0.144)	(0.164)
Government					1.696	2.383
					(1.864)	(2.356)
Non-Profit					0.818	0.825
					(0.730)	(0.755)
Residential					−1.497***	−1.548***
					(0.434)	(0.433)
N	67,230	67,230	67,230	67,230	67,230	67,230
1st-stage partial F-stat.	-	71982.4	-	47312.9; 10032.1	-	319.4; 32.4; 19.6; 24239.2

Notes: The outcome variable is the design factor of TPO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters. The 1st stage partial F-statistics for the instrumental variables are derived from first- stage regression results reported in Appendix B, Table B1. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

5 Empirical results

I start this section by presenting the impact of subsidy levels on the PV system's design factor (Section 5.1) followed by transaction prices per Watt (Section 5.2). I further analyze implications from the framework in Section 2 and study heterogeneity related to increased verification measures and the ownership of PV systems.

5.1 Subsidy levels and the design factor of upfront systems

Table 3 shows regression results for Equation 1 when the outcome is the design factor of TPO upfront systems and all observations are included.[10] In column (1), I report OLS

---

[10] Throughout this section, I use the Stata package REGHDFE to estimate linear models with multiple fixed effects (Correia, 2019). I exclude singleton groups (i.e. groups with only one observation) to avoid underestimated standard errors which could bias statistical inference (Correia, 2015). Keeping singleton groups does not affect the qualitative conclusions.

estimates and in column (2), I report 2-stage least squares (2SLS) estimates where the actually received subsidy level is instrumented with the predetermined subsidy level as shown in equation 2. In columns (3) and (4) I repeat this sequence and interact the subsidy level with a variable equal to one if the system is subject to a mandatory field inspection. In columns (5) and (6) I interact the subsidy level with a variable indicating the sector of the consumer. Table C1 in Appendix C, shows regression results for the same sequence dropping observations in the vicinity of a subsidy level drop date.

In columns (1), and (2), the coefficient on *Subsidy level* is positive and statistically significant. When interacting the subsidy level with a variable indicating whether the system is subject to a mandatory field inspection (*FI = 1*) in columns (3) and (4) the positive association of subsidy levels and the design factor of systems which are not subject to a mandatory field inspection (*Subsidy level*) is similar in size and significance. At the same time, there is no statistically significant association of subsidy levels and the design factor when systems are subject to a mandatory field inspection (*FI = 1 x Subsidy level*). Looking at the marginal effect of the subsidy level by consumer sector in columns (5) and (6), we observe that only residential consumers show statistically significant and positive coefficients which are again similar in size to the coefficients without interaction. Further, residential systems are associated with a statistically significant lower design factor of approximately 1.5 percentage points.

The estimates after dropping observations in the vicinity of a rebate level drop date in Table C1, Appendix C are very similar in size and significance, suggesting that the presence of consumers anticipating such dates does not affect the association of subsidy levels and the design factor. Furthermore, the OLS and IV estimates are similar in size, although the IV estimates tend to be larger. A negative endogeneity bias suggests that any omitted variable influencing both the subsidy level and the error term lowers the association between subsidy levels and the design factor. This indicates that the received subsidy levels were not influenced by factors also increasing the design factor (such as for example second-degree moral hazard). Using the predetermined subsidy level as an instrument for the actually received subsidy level further has significant explanatory power indicated by large first-stage F-statistics.

The results show that a one dollar increase of the upfront subsidy is associated with an increase of the design factor of approximately 0.47 percentage points (in my preferred specification in column 2). Following sample calculations of the design factor, an increase of 0.47 percentage points is equivalent to an optimization of the module tilt of 5 degrees.[11] In combination with the observation that there is no such association when systems are subject to a mandatory field inspection where the inputs to calculate the design factor are verified, I interpret these results as evidence that TPO installers do respond to larger subsidy levels and increase the design factor of systems, unless their input is verified during mandatory field inspections.

The positive association between subsidy levels and the design factor for residential consumers is further in line with heterogeneous social preferences depending on the consumer sector as argued in section 2. Inflating and shifting the total amount of subsidies to consumers has the consequence that consumers pay less for their systems. Lower design factors of residential consumers can further potentially be attributed to different intentions to invest in a PV system. Investments by residential consumers may be motivated by an environmental perspective, which implies that the investments are conducted although the location of their system may not be optimal to generate electricity output and therefore have a lower design factor, such as carport structures. Instead, commercial consumers may only want to invest if the location of the system has high potential for large electricity output.

Next, Table 4 shows regression results for Equation 1 when the outcome is the design factor of HO upfront systems. The columns are arranged in the same way as in Table 3. Table C1 in Appendix C, shows regression results for the same sequence dropping observations in the vicinity of a subsidy level drop date.

While there is no statistically significant association of subsidy levels and the design factor, mandatory field inspections are associated with a significant decrease of the design factor of 0.43 percentage points and the design factor of residential systems is approximately 0.9 percentage lower compared to commercial systems. The estimates after dropping observations in vicinity of a subsidy level drop date shown in Table C1, Appendix C are similar in size and significance.

These results suggest that HO systems do not have an increased design factor when they receive larger subsidy levels. The observation that HO systems subject to a mandatory field inspection are associated with a significantly lower design factor suggests that increased verification may trigger conservative reports of system characteristics, as any discrepancies are more likely to be detected. Furthermore, residential consumers are again associated with a lower design factor indicating that residential consumers install PV systems in areas with lower potential to generate electricity output.

---

[11] See http://www.csi-epbb.com/ for further information and sample calculations.

Table 4: Design factor of HO systems

	All obs. included					
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Subsidy level	0.010	−0.067	0.006	−0.071	−0.015	−0.135
	(0.105)	(0.122)	(0.105)	(0.122)	(0.206)	(0.255)
Field inspection (FI)						
FI = 1 × Subsidy level			0.117	0.074		
			(0.163)	(0.176)		
FI = 1			−0.410**	−0.434**		
			(0.145)	(0.149)		
Sector						
Government × Subsidy level					−0.084	−0.436
					(0.549)	(0.662)
Non-Profit × Subsidy level					−0.451	−0.467
					(0.307)	(0.348)
Residential × Subsidy level					0.005	−0.064
					(0.105)	(0.121)
Government					−0.551	−0.208
					(0.704)	(0.849)
Non-Profit					0.623	0.617
					(0.383)	(0.421)
Residential					−0.923***	−0.913***
					(0.145)	(0.145)
N	69,113	69,113	69,113	69,113	69,113	69,113
1st-stage partial F-stat.	-	43892.7	-	88043.9; 34064.67	-	3090.8; 296.7; 1018.9; 1.2e+5

Notes: The outcome variable is the design factor of HO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters. The 1st stage partial F-statistics for the instrumental variables are derived from first- stage regression results reported in Appendix B, Table B1. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

Overall, these results suggest that TPO installers increase the design factor of PV systems when they receive higher subsidy levels. This association indicates second-degree moral hazard which, however, does not come at the expense of the consumer but at that of the CSI as the subsidy provider. Because TPO installers receive the total amount of subsidies directly after the installation has been completed and the transaction price paid by the consumer is paid dispersed over time, increasing the amount of subsidies can increase the short-term cash flow. Instead, installers of HO systems receive the full transaction price and they do not receive the subsidies as they are directed to the consumer. Increasing short-term cash flow for HO installations would hence lead to larger transaction prices, which I study in the next section.

5.2 Subsidy levels and transaction prices per Watt of PV systems

In this section, I study the association of subsidy levels and transaction prices per Watt of PV systems and how it depends on increased verification and the consumer sector. I first analyze the transaction price of TPO systems (which has been acknowledged to

Table 5: Transaction price per Watt of TPO systems

	All obs. included					
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Subsidy level	0.045	0.001	0.045	0.001	−0.075	−0.218
	(0.031)	(0.030)	(0.031)	(0.030)	(0.173)	(0.206)
Field inspection (FI)						
FI = 1 × Subsidy level			0.059	0.006		
			(0.149)	(0.150)		
FI = 1			−0.094	−0.091		
			(0.109)	(0.109)		
Sector						
Government × Subsidy level					2.477^{**}	2.358^{*}
					(0.937)	(0.945)
Non-Profit × Subsidy level					0.208	0.297
					(0.277)	(0.490)
Residential × Subsidy level					0.042	0.002
					(0.031)	(0.030)
Government					−0.243	−0.213
					(0.542)	(0.620)
Non-Profit					0.203	0.223
					(0.212)	(0.223)
Residential					0.243^{*}	0.254^{*}
					(0.120)	(0.121)
N	67,230	67,230	67,230	67,230	67,230	67,230
1st-stage partial F-stat.	-	72154.6	-	47481.7; 9919.2	-	313.2; 33.0; 19.59; 24436.8

Notes: The outcome variable is the transaction price per Watt of TPO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters, the experience of installers, the relative market power of installers and a measure for local industry concentration. The 1st stage partial F-statistics for the instrumental variables are derived from first-stage regression results reported in Appendix B, Table B2. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

be inconsistently reported and results are therefore indicative) and then redo the same analysis for HO systems.

Table 5 shows regression results for Equation 1 when the outcome is the transaction price per Watt of upfront TPO systems. The columns are arranged in the same way as in the previous regression Tables. Table C3 in Appendix C, shows regression results for the same sequence dropping observations in the vicinity of a subsidy level drop date. I recall the disclaimer, that all qualitative and quantitative conclusions related to the transaction prices of TPO systems are indicative, because transaction prices reported to the CSI may not necessarily represent what the consumer is actually paying to the TPO installer (see Section 3.5 for further details).

Column (6) shows that there is a large, positive and statistically significant association of the marginal subsidy level and transaction prices for governmental consumers. A one dollar increase of the subsidy level is associated with a 2.4 $ increase of the transaction price per Watt. Given average transaction prices per Watt range between 4 and 8 $

Table 6: Transaction price per Watt of HO systems

	All obs. included					
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Subsidy level	0.234^{***}	0.298^{***}	0.233^{***}	0.297^{***}	0.384^{***}	0.402^{***}
	(0.035)	(0.047)	(0.035)	(0.047)	(0.090)	(0.102)
Field inspection (FI)						
FI = 1 × Subsidy level			0.247^{***}	0.312^{***}		
			(0.064)	(0.064)		
FI = 1			-0.095^{**}	-0.095^{*}		
			(0.037)	(0.037)		
Sector						
Government × Subsidy level					0.645^{**}	0.598^{*}
					(0.220)	(0.263)
Non-Profit × Subsidy level					0.649^{***}	0.715^{***}
					(0.107)	(0.119)
Residential × Subsidy level					0.245^{***}	0.295^{***}
					(0.036)	(0.047)
Government					0.440	0.506
					(0.260)	(0.314)
Non- Profit					-0.537^{***}	-0.590^{***}
					(0.114)	(0.127)
Residential					-0.137^{**}	-0.151^{**}
					(0.052)	(0.051)
N	69,113	69,113	69,113	69,113	69,113	69,113
1st-stage partial F-stat.	-	43965.8	-	87944.67; 33857.3	-	3102.2; 296.8; 1018.5; 1.2e+5

Notes: The outcome variable is the transaction price per Watt of HO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters, the experience of installers, the relative market power of installers and a measure for local industry concentration. The 1st stage partial F-statistics for the instrumental variables are derived from first- stage regression results reported in Appendix B, Table B2. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

per Watt (see Table 2), this translates to 50 to 25 percent increase of the transaction price per Watt. This observation is in line with installers having different valuations for consumer types (i.e. with heterogeneity of λ in equation 2). The transaction prices of governmental consumers are ultimately paid by the tax-payer, which adds another layer of third party reimbursements and may reduce the extent to which installers care for the consumer's benefits λ leading to larger transaction prices.

Furthermore, residential consumers are associated with a larger transaction price per Watt. This association is however not statistically significant when dropping observations in the vicinity of a rebate level drop date (Table C3, Appendix C). A larger transaction price per Watt could again be attributed to less standardized installations in residential consumers, such as carport structures (see Gillingham et al., 2016a, for a similar reasoning).

Next, Table 6 shows regression results for Equation 1 when the outcome is the transaction price per Watt of HO systems. The columns are arranged in the same way as in the previous regression Tables. Table C4 in Appendix C, shows regression results for the same sequence dropping observations in the vicinity of a subsidy level drop date.

The association of subsidy levels and the transaction price per Watt is positive and statistically significant in columns (1) and (2). When interacting the subsidy level with a variable indicating whether the system is subject to a mandatory field inspection (*FI = 1*) in columns (3) and (4), we do not observe a significant difference of the marginal effect of the subsidy level with respect to field inspections (Wald test column 4, $H_0 =$ *Subsidy level = FI=1 x Subsidy level*, $p = 0.786$). At the same time, a mandatory field inspection is further associated with a statistically significant decrease of the transaction price per Watt of approximately 0.1 $ per Watt.

The marginal effects of the subsidy level on transaction prices per Watt by the consumer sector in columns (5) and (6) are all positive and statistically significant. Pairwise comparisons of the coefficients show that the marginal effect for residential consumers is significantly lower than that of non-profit consumers (Wald test column 6, $H_0 =$ *Non Profit x Subsidy level = Residential x Subsidy level*, $p < 0.001$). Finally, residential and non-profit systems are associated with a statistically significant decrease of the transaction price per Watt compared to commercial consumers.

These results suggest that a one dollar increase of upfront subsidies is associated with an increase of the transaction price per Watt of approximately $ 0.3 per Watt. Given average transaction prices per Watt range between 4 and 8 $ per Watt (see Table 2), this translates to a 7.5 to 3.7 percent increase of the transaction price per Watt. While mandatory field inspections do not reduce this association, they are associated with lower transaction prices per Watt, suggesting that they may trigger conservative pricing of installers as any surcharges are more likely to be detected.

The observation that the marginal effect of subsidy levels on transaction prices per Watt for non-profit consumers is the largest, is perhaps striking because non-profit organizations may be perceived as serving a good cause with little financial resources. There have, however, been some discussions on a decreased confidence in non-profit organizations as the sector is mostly unregulated and sometimes viewed as unethical as their good cause and "non-profit" status is doubted (O'Neill, 2009). This may in turn reduce the extent in how far installers care for the NPO's benefits λ and therefore lead to higher transaction prices.

In combination with the insights from Section 5.1, the results suggest that TPO installers do increase the design factor of residential consumers and at the same time only increase the transaction prices per Watt for governmental consumers. For HO systems, only the transaction price per Watt and not the design factor is increasing in the subsidy level. These insights are in line with increasing the short-term cash flow of installers, because directly after the installation TPO installers receive the total subsidies and HO installers receive the transaction price. In addition, the results can inform on the subsidy pass-through of TPO installers as they increase the total amount of subsidies for residential consumers but do not adapt transaction prices accordingly which would ultimately lead to a larger pass-through.

6 Discussion and conclusion

In this paper, I studied second-degree moral hazard of installers induced by the credence component of energy-transforming technologies. To this end, I analyzed data from a solar subsidy program in California and quantified the relationship of subsidy levels and the design factor as well as transaction prices per Watt of PV systems. Employing an instrumental strategy to account for potential self-selection of installers into specific subsidy levels and further controlling for a wide range of potential confounding factors such as an installer's market power, I find that TPO installers increase the design factor and thereby the total amount of subsidies received for residential systems. Such second-degree moral hazard at expense of the subsidy provider is however non-existent when the system is subject to a mandatory field inspection.

TPO installers do further not adapt transaction prices per Watt when residential consumers receive larger subsidies. In addition, the design factor of HO systems is unaffected by larger subsidy levels, but I find evidence suggesting that second-degree moral hazard increases the transaction prices per Watt paid by all consumer sectors. This is in line with the intention of installers to increase their short-term cash flows.

The results provide novel insights for different kinds of research avenues. First, they contribute to the literature evaluating the cost-effectiveness of environmental subsidy programs and show that such programs need to be robust towards second-degree moral hazard of installers induced by the credence component of energy-transforming technologies. My empirical analysis suggests that program administrators should i) account for the cost of second-degree moral hazard when installers can to some extent determine the total amount of subsidies received and ii) impose stricter verification measures on the work of installers.

Second, I document that stylized findings from other credence goods markets are relevant for the market of energy-transforming technologies. This paper is the first to document evidence of second-degree moral hazard in the context of energy-transforming technologies. The results further confirm heterogeneity of second-degree moral hazard depending on increased verification and the bearer of the financial consequences (Balafoutas et al., 2013; Gottschalk et al., 2020). It is a promising route for future research to further uncover dimensions of heterogeneity in second-degree moral hazard as well as related solutions.

Chapter III

Reciprocity and gift exchange in markets for credence goods

This chapter is mainly based on a paper co-authored by Bruno Lanz (University of Neucht̂el) and Serhiy Kandul (University of Zürich). This study is registered in the AEA RCT Registry under identification number AEARCTR-0004213. This paper has been published in the working paper series of UNINE IRENE (No. 20-09). It was presented by Evert Reins at the Annual Conference of the Association for Research in Economic Psychology (online), 2021, the Annual Conference of the Society for the Advancement of Behavioral Economics (online), 2020 and the 6th CUSO workshop in Experimental Research, Neuchâtel (Switzerland), 2018. It has further been accepted to be presented at the Annual Congress of the Swiss Society of Economics and Statistics (SSES) in Zürich (Switzerland), 2020 (canceled).

Abstract

We study the role of reciprocity in markets where expert sellers have more information about the severity of a problem faced by a consumer. We employ a standard experimental credence goods market to introduce the possibility for consumers to gift the expert seller before the diagnostic, where the gift is either transferred unconditionally or conditionally on solving the problem. We find that both types of gift reduce undertreatment, whereas unconditional gifts also reduce overcharging and increase undercharging, suggesting that unconditional gifts are perceived as more kind. For high-severity consumers gifting reduces market inefficiencies, although the presence of low-severity consumers mitigates overall efficiency gains.

Keywords Credence Goods; Gift Exchange; Asymmetric Information; Lab Experiment.
JEL codes D82; L14; C91.

1 Introduction

Many people do not know how to repair a broken heater and therefore ask a repairman to diagnose and fix the problem. Since the repairman knows more about the severity of the problem, he may have an incentive to provide a service that maximizes own profits instead of meeting the consumer's needs (Emons, 1997; Dulleck and Kerschbamer, 2006). Knowing this, the consumer may naturally consider offering a cup of coffee to the repairman, hoping to establish a reciprocal relationship and secure consumer-friendly actions. This intuition is supported by a large literature showing that gifting by a principal increases the efforts of an agent, thereby also increasing the principal's profit.[1] In markets for goods with a credence component, however, consumers are able to observe if their problem is solved, but they cannot verify that the service provided and the price charged are adequate. In turn, actions by the agent are partially hidden from the principal, which reduces the scope for reciprocity (Güth et al., 1996; Andreoni and Bernheim, 2009; Hoppe and Schmitz, 2018).

In this paper we provide experimentally controlled evidence on how gift exchange and reciprocal expert sellers (e.g. Falk and Fischbacher, 2006) affect inefficiencies on markets for credence goods.[2] We employ the experimental framework of Dulleck et al. (2011) in which a consumer faces a problem of either high or low severity and needs the corresponding high- or low-quality service to fix it. After observing the price for each service, the consumer may decide to interact with the expert seller. In this case, the expert seller learns which service is needed by the consumer (akin to a diagnostic), supplies one of the two services, and subsequently charges one of the two prices independently of the service actually provided.

In the baseline condition (BASE), which allows us to document the behavior of consumers and expert sellers without the possibility to gift, the parametrization of the experiment implies that expert sellers have an incentive to provide the low-quality service to consumers in need of a high-quality service (undertreatment) and charge for the high-quality service (overcharging). In turn, consumers are better off not interacting with expert sellers. The standard prediction therefore implies that all consumers opt out of the market, leading to market collapse (akin to Akerlof, 1970).

---

[1] Examples notably include paying more than the market wages to increase workers' efforts (Akerlof, 1982; Fehr et al., 1993, 1998; Abeler et al., 2010; Kube et al., 2012; Cohn et al., 2015) and granting small gifts to increase charity donations by potential donors (Falk, 2007; Carpenter, 2017).

[2] Related studies have shown that consumer-friendly actions are more likely to emerge in the presence of expert sellers who are guilt averse (Beck et al., 2013), hold altruistic preferences (Hennig-Schmidt et al., 2011; Godager and Wiesen, 2013), or are inequality averse (Kerschbamer et al., 2017).

In the gift exchange (GE) treatment, we extend BASE by giving consumers the possibility to gift the expert seller before the diagnostic takes place. More specifically, consumers can transfer part of their payoff as a bonus payment to the expert seller. Importantly, we consider a "small" gift equal to the smallest integer of our experimental currency (see Malmendier and Schmidt, 2017, for a similar procedure).[3] Based on the reciprocity model of Falk and Fischbacher (2006), we first show that expert sellers are expected to perceive the transfer as a kind action. In turn, gifting can induce expert sellers with a preference for reciprocity to abstain from undertreatment and overcharging. More interestingly, when the consumer is of the high-severity type, expert sellers can also reciprocate a gift by supplying the high-quality service and charging for the low-quality service. The possibility to *undercharge* is akin to offering a discount on performing the high-quality service and constitutes the strongest form of reciprocity in our context. Importantly, while consumers can observe whether the problem has been solved, they cannot verify the type of service provided and therefore reciprocal actions by the expert sellers are not observed by the consumer.

Next, we investigate the effects of a conditional gift (GEC treatment), whereby the consumer commits to sending a gift before the diagnostic and the gift is transferred only if the expert seller supplies a service of sufficient quality. As in the GE treatment, the gift does not change the payoff-maximizing behavior of expert sellers. However, it partially aligns incentives and is akin to a form of contracting over the gains from a sufficient treatment (see Bester and Dahm, 2017). At the same time, because the conditional gift imposes a minimum performance level on the expert seller, it can be perceived as a sign of distrust (see e.g. Fehr and List, 2004; Falk and Kosfeld, 2006). In turn, if the conditional gift is perceived as less kind than an unconditional gift, the framework by Falk and Fischbacher (2006) suggests that the reciprocal response by expert sellers in the GEC treatment will be lower.

Results from our experiment show that consumers who interact with expert sellers gift in 35% of interactions in the GE treatment and 45% in the GEC treatment, and that gifting induces expert sellers to perform reciprocal actions even if those are not observed by the consumer. More specifically, the likelihood of undertreatment declines by about 18 percentage points following an unconditional gift, and by around 14 percent after a conditional gift. In addition, an unconditional gift increases undercharging by around

---

[3] While small, the gift represents ten percent of the surplus associated with solving the problem of the consumer and is therefore not symbolic. Evidence suggests, however, that both material and immaterial gifts trigger reciprocal behavior (see e.g. Kirchler and Palan, 2018, in the context of experience goods).

18 percentage points and reduces overcharging by around eight percentage points, whereas undercharging and overcharging are not significantly affected by conditional gifts. This is consistent with the presence of expert sellers with preferences for reciprocity, and suggests that a conditional gift is perceived as less kind than an unconditional gift.

In line with this, we find that consumers who face a high-severity problem and gift the expert seller earn on average higher profits. In turn, for these consumers gift exchange results in higher market efficiency (a measure of total profits by expert sellers and consumers, see Dulleck et al., 2011). For consumers with a low-severity problem, gifting either does not produce efficiency gains (conditional gifts) or produces efficiency losses (unconditional gifts). In the latter case, we show that these can be attributed to an increase in overtreatment, which has been interpreted as a manifestation of preferences for equal payoffs (Kerschbamer et al., 2017).[4] Taken together, our results suggest that gift exchange may improve market outcomes when the share of consumers with high-severity problems is large, or when agents have private information about own type.

This paper contributes to a growing literature that investigates how different characteristics of credence good markets affect the behavior of expert sellers (see Balafoutas and Kerschbamer, 2020, for a recent overview). Examples include imposing liability and/or verifiability (Dulleck et al., 2011; Mimra et al., 2016a), enhancing competition and reputational concerns (Rasch and Waibel, 2017; Soraperra et al., 2019), manipulating the information available to consumers (Balafoutas et al., 2013; Agarwal et al., 2019; Mimra et al., 2016b), insurances and third party reimbursement (Kerschbamer et al., 2016; Huck et al., 2016; Balafoutas et al., 2017), or introducing non-binding promises (Beck et al., 2013). Related to our study, Kerschbamer et al. (2017) use the experimental credence goods market of Dulleck et al. (2011) to show that less than a fourth of expert sellers conform with canonical preferences for own material payoffs. Instead, there is significant heterogeneity among expert sellers, with a majority displaying some form of aversion to inequality (as in Andreoni and Miller, 2002; Charness and Rabin, 2002; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Kerschbamer et al. (2017) also highlight that consumers face the tedious task of identifying prosocial expert sellers so as to receive more consumer-friendly services, and our results suggest that gifting can potentially replace complicated selection

---

[4] In our experiment, the expert seller has the possibility to overtreat consumers who face a low-severity problem by supplying the high-quality treatment. However, given the standard parametrization of the experiment, overtreatment is strictly dominated by overcharging. We come back to this below.

mechanisms.

Our work is also closely related to the field study of Currie et al. (2013), which focuses on a specific patient-physician setting in China and shows that physicians who receive a “token” gift (a self-made bookmark) spend more time with gift-giving patients and prescribe fewer unnecessary drugs (see also Currie et al., 2014). Our experimentally controlled results disentangle the impact of gifting across possible actions by expert sellers and provide novel evidence on market efficiency impacts. To the best of our knowledge, this paper is also the first to document the possibility of undercharging in an experimental credence goods market, suggesting that unconditional gift can also induce expert sellers to offer a form of discount.

Our results also contribute to a wider literature on gift exchange. First, our findings are in line with principal-agent experiments focusing on bonus payments (see for example Angelova and Regner, 2018; Soraperra et al., 2019). In these studies, bonus payments made after observing an agent’s decisions change the agent’s behavior such that it results in higher payoffs for principals. Relative to these studies, we show that gifts have a positive effect even when the agent has no reputational concerns and the principal cannot observe the behavior of the agent. Second, our work is related to studies showing that observability matters for reciprocity (see Bradley et al., 2018, for a review). For example, using the principal-agent game of Charness and Dufwenberg (2006), Hoppe and Schmitz (2018) report a large drop in reciprocity when the agent’s action becomes unobservable (see also Rubin and Sheremeta, 2015; Davis et al., 2017). In our work, we show that expert sellers reciprocate unconditional gifts with a range of unobservable actions, while conditional gifts only reduce undertreatment. This finding is also related to Falk and Kosfeld (2006) who show that payments conditioned on a minimum performance reduce effort, and Newman and Jeremy Shen (2012) where charitable donations decline when a small material gift is offered conditional on a positive donation.

The remainder of this paper is organized as follows. Section 2 lays out the experimental design. In Section 3 we use the framework by Falk and Fischbacher (2006) to derive our main hypotheses. We present our results in Section 4. Section 5 concludes.

2 Experimental Design

This section first presents the experimental credence goods market of Dulleck et al. (2011), which represents the baseline treatment in our study. We then introduce two experimental treatments in which the consumer is given the possibility to gift the expert seller, either unconditionally (GE treatment) or conditionally on receiving a service of sufficient quality (GEC treatment). Lastly, we provide details about implementation and data collection.

2.1 Baseline experimental credence goods market (BASE treatment)

Consider a consumer with a problem that is of either high or low severity. The consumer, however, only knows that a high-quality service q_h is needed with probability h and a low-quality service q_l is needed with probability $(1 - h)$, where $h = 0.5$.[5] The expert seller can provide q_h, which solves both the high- and low-severity problems, at cost $c_h = 6$. Alternatively, supplying q_l only solves the low-severity problem ($c_l = 2$). Both c_h and c_l are known by consumers.

The extensive form of the game in BASE is depicted in Figure 1. The game comprises four decisions: decisions 1, 3 and 4 are made by the expert seller, decision 2 is made by the consumer. At decision 1, the expert seller announces prices p_h and p_l. Both prices must be integers between 1 and 11, with $p_h \geq p_l$.[6] At decision 2, the consumer observes p_h and p_l, and decides whether to interact with the expert seller. If the consumer opts out of the market, the game stops and both players receive the outside option $o = 1.6$. If the consumer opts in, the game moves on to a third stage in which the expert seller learns about the severity of the consumer's problem (diagnostic stage). Based on this, in decision 3 the expert seller supplies either q_h or q_l, and in decision 4 either p_h or p_l is charged. Importantly, the expert seller can charge p_h or p_l independently of the service provided, and the consumer is not able to verify whether q_h or q_l is supplied.

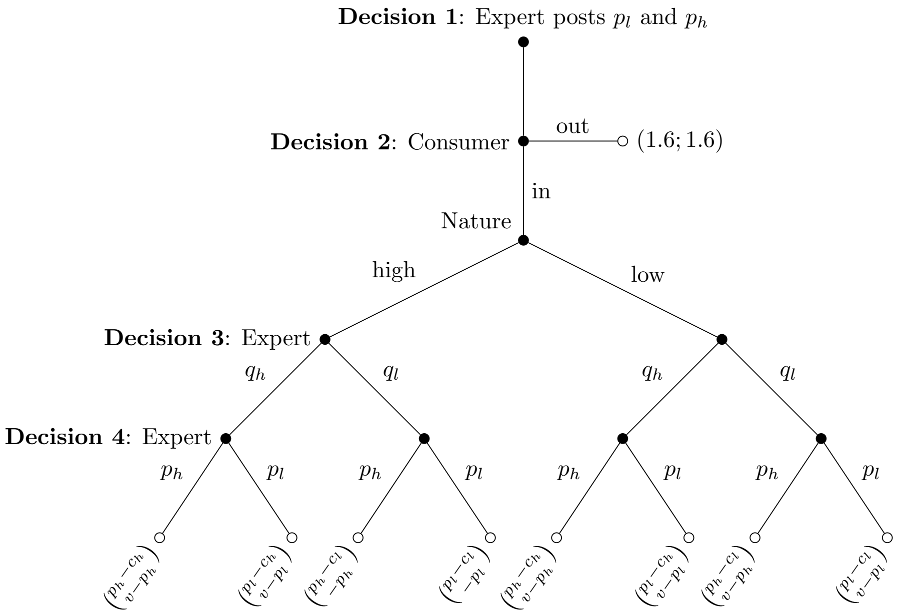
At the end of the game, the payoffs are determined as follows. If the problem is solved (i.e. the consumer needs q_l and receives either q_l or q_h, or the consumer needs q_h and receives q_h), the consumer receives $v = 10$ points and pays the price charged by the expert seller. The payoff of the consumer is therefore: $\pi_c = v - p_i$ ($i \in \{h, l\}$). If the problem is not solved (the consumer needs q_h but receives q_l), $v = 0$ and hence $\pi_c = -p_i$. One implication is that consumers observe when they have been undertreated, whereas

---

[5] All the parameters we use in the experiment are identical to those in the baseline treatment (B/N) of Dulleck et al. (2011).

[6] Related laboratory experiments set prices exogenously, thereby creating incentives for particular supply side inefficiencies (e.g. overtreatment in Mimra et al., 2016b; Huck et al., 2016). Instead, we retain the original procedure of Dulleck et al. (2011) to provide a general account of market inefficiencies in this context before introducing the possibility for gift exchange.

Figure 1: Extensive form game of the BASE treatment



Notes: Payoffs are shown in vectors at the end nodes. The first row of the payoff vector denotes the expert seller's profit, the second row is the consumer's profit.

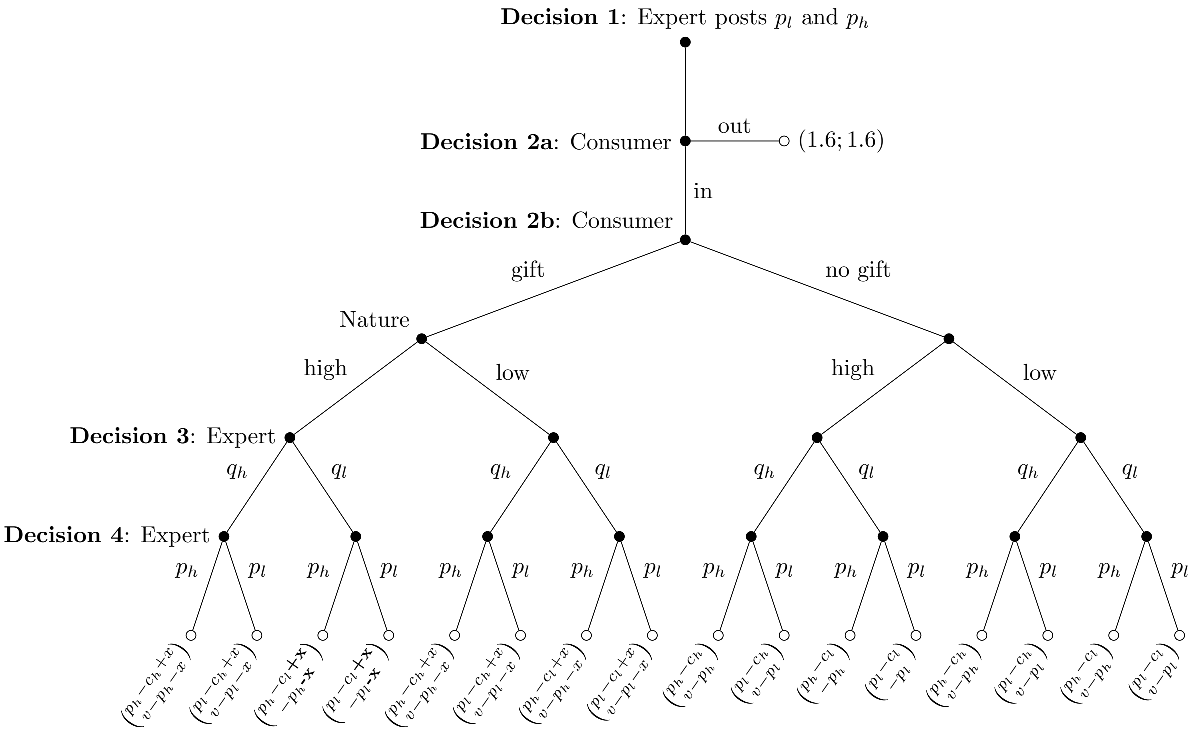
they do not know if they have been overcharged, undercharged or overtreated. The payoff of the expert seller is simply the difference between the price charged and the cost of the treatment supplied: $\pi_e = p_i - c_i$.

2.2 Unconditional gift exchange (GE treatment)

This treatment extends BASE by giving consumers the possibility to unconditionally gift the expert seller before the diagnostic. As shown in Figure 2, after the decision to interact with the expert seller (decision 2a), the consumer can transfer $x \in \{0; 1\}$ to the expert seller (decision 2b). The expert seller is then informed about whether or not the consumer has decided to send a gift, learns about the problem faced by the consumer, and selects the service performed (decision 3) and the price charged (decision 4). Accordingly the payoff for the consumer is $\pi_c = v - p_i - x$ if the problem is solved and $\pi_c = -p_i - x$ if it is not, and the expert seller receives $\pi_e = p_i - c_i + x$.

As mentioned above, the objective of this treatment is to study the effects of a small gift, and we therefore exogenously set the size of the gift to the smallest integer unit $x = 1$

Figure 2: Extensive form game of the GE and GEC treatments



Notes: Payoffs are shown in vectors at the end nodes. The first row of the payoff vector denotes the expert seller's profit, the second row is the consumer's profit. Payoffs reported in end nodes three and four are marked in bold because the transfer is only realized in the GE treatment, not in the GEC treatment.

(as in Malmendier and Schmidt, 2017).[7] Moreover, as the gift represents a transfer from the consumer to the expert seller, the gift has no direct impact on total market surplus. In turn, this mitigates efficiency-seeking motives for a consumer to send the gift.

2.3 Conditional gift exchange (GEC treatment)

The GEC treatment is identical to the GE treatment except that the transfer is realized only if the expert seller solves the consumer's problem. More specifically, after having decided to interact with the expert seller in decision 2a, in decision 2b the consumer commits to a transfer of $x = 1$ if q_l is needed (either q_l or q_h can be provided) or if q_h is needed and the expert seller provides q_h (i.e. no undertreatment). As shown in the extensive form game (Figure 2), only the payoffs in the third and fourth end nodes are affected. Note that conditioning the gift on the provision of sufficient quality is possible because the consumer observes whether the problem is solved or not.

A conditional gift partially aligns the incentives of the expert seller and the consumer and can be interpreted as a form of contracting where an expert seller who performs a service of sufficient quality is entitled to a share of the surplus.[8] Bester and Dahm (2017) for example argue that physicians could be paid conditionally on the patient's satisfaction and further show that contracting generally increases efficiency in markets for credence goods. However, conditioning the transfer of the gift on a minimum performance requirement might backfire because the expert seller could understand it as a sign of distrust (Fehr and List, 2004; Falk and Kosfeld, 2006).[9] As we discuss below, comparing GE and GEC treatments can therefore provide evidence about the reciprocity motive underlying the behavior of the expert seller.

2.4 Experimental Procedure

The experiment was run in the laboratory of the University of Neuchâtel in October 2019 and implemented in z-Tree (Fischbacher, 2007). We recruited a total of 168

---

[7] As we discuss below, fixing the size of the gift to $x = 1$ ensures that the payoff maximizing strategies are not altered. However, evidence from the literature suggests that the intentions behind gift-giving matter more than the size of the gift (Hannan et al., 2002; Newman and Jeremy Shen, 2012; Kube et al., 2012), and this design choice is unlikely to affect our conclusions.

[8] The size of the gift $x = 1$ ensures that the conditional gift does not completely align the incentives of the expert seller and the consumer, as undertreatment still increases the profits of the expert seller.

[9] The conditional gift also provides the expert seller with monetary incentives to abstain from undertreatment. Monetary incentives have shown to crowd out intrinsic motivation to fulfill a task (Frey and Oberholzer-Gee, 1997; Gneezy and Rustichini, 2000; Mellström and Johannesson, 2008; Chao, 2017) which could ultimately lead the expert seller to provide less consumer-friendly services.

participants via invitation emails sent to all students which were allocated equally to the three experimental treatments. There were four experimental sessions per treatment, out of which 3 sessions were conducted with 16 participants and one session was conducted with 8.[10]

The following relevant procedural factors were adopted from Dulleck et al. (2011). The framing of the instructions was neutral, we did for example not talk about expert sellers and consumers but about "role A" and "role B." Participants were randomly assigned to one of the roles at the beginning of the experiment and stayed in that role throughout the experiment. Matching groups of eight subjects were randomly formed at the beginning of the experiment, bringing together four consumers and four expert sellers. [11] The stage game in each treatment (see Figures 1 and 2) was repeated for 16 periods, and each consumer was randomly matched with one expert seller at the beginning of each period.[12]

Upon arrival, each participant was randomly allocated to a cabin and started reading the instructions which were also read aloud 10 minutes after all participants were seated. Before the stage game started for the first time, participants had to correctly answer a set of control questions. In the first period, each participant received an initial endowment of 6 points. The participant's earnings were summed up over the 16 periods and then converted at an exchange rate of 2 points = 1 CHF ($\approx$ US\$ 1). Together with a show up fee of CHF 10, participants earned on average about CHF 30 and sessions lasted approximately 80 minutes.

3 Theoretical predictions and hypotheses

This section discusses predictions for the stage games shown in Figures 1 and 2. We first describe standard predictions for self-interested players. We then use the general theory of reciprocity by Falk and Fischbacher (2006) to derive implications of introducing gift exchange in the experimental market for credence goods.

---

[10] For each session we invited more participants than required and once the targeted number was reached the remaining participants were paid a show up fee of CHF 10 ($\approx$ US\$ 10) and dismissed.

[11] Thus our experiment includes seven matching groups per treatment with eight participants in each, which is comparable to other applications of the experimental credence goods markets, such as Dulleck et al. (2011), between six and 12 matching groups, seven in Huck et al. (2016), and eight in Mimra et al. (2016a) and Beck et al. (2014).

[12] We employed a stranger matching protocol to avoid reputational concerns. Over the course of the game, each consumer interacted with each expert seller four times but could not know in which period it would happen.

3.1 Self-interested expert sellers

Standard predictions for the experimental credence goods markets are derived from the equilibrium characterized in Dulleck et al. (2011) and are based on self-interested agents who maximize own payoffs.[13] This implies that expert sellers always supply the low-quality service q_l and charges for the high-quality one p_h. Moreover, expert sellers always post prices such that $\pi_e = p_h - c_l \geq o$, which implies $p_h \geq 4$ since only integers are allowed.

The consumer therefore anticipates undertreatment if q_h is needed and overcharging if q_l is needed, so that his expected payoff is $\pi_c = h \cdot (-p_h) + (1-h) \cdot (v - p_h)$. Given expectations about prices, the payoff from interacting with an expert seller is strictly lower than the outside option ($\pi_c < o$), and it is optimal for consumers to stay out of the market. In turn, the standard prediction implies that the market in BASE collapses.

The possibility to receive a gift does not affect the payoff maximizing strategy (q_l, p_h) of the expert seller (since $c_h - c_l > 1$). In both GE and GEC, it is therefore always optimal to undertreat consumers even if it implies not receiving the conditional gift.[14] For the consumer, this implies that (i) sending a gift always decreases the expected payoff and (ii) opting out of the market is the payoff maximizing strategy. In turn, the standard prediction also implies market collapse in both GE and GEC.

3.2 Reciprocal expert sellers

The predictions change considerably if expert sellers have a disposition for reciprocity and are willing to sacrifice part of their material payoff to reciprocate a kind action of the consumer. Formally, we follow Falk and Fischbacher (2006) and write the utility function of a reciprocal expert seller e as:

$$U_e(a_e, a_c) = \underbrace{\pi_e(a_e, a_c)}_{\text{material payoff}} + \underbrace{\rho_e \cdot \phi_c(a_c) \cdot \sigma_e(a_e)}_{\text{reciprocity utility}} \tag{1}$$

where both the material payoff and reciprocity utility depend on the actions of the expert seller a_e and those of the consumer a_c. In this framework, reciprocity utility is driven by three parameters: the reciprocity parameter ρ_e, the kindness term $\phi_c(a_c)$ and

---

[13] This equilibrium assumes that agents play each of the 16 rounds as a one-shot interaction, which is consistent with random re-matching in every period. See Dulleck et al. (2011) for a discussion of reputation equilibria.

[14] In the GEC treatment, if a gift-giving consumer needs q_h, the profit-maximizing strategy (q_l, p_h) yields $\pi_e = p_h - c_l$ whereas playing (q_h, p_h) yields $\pi_e = p_h - c_h + 1$.

the reciprocation term $\sigma_e(a_e)$. We now discuss these in turn.

The first component, $\rho_e \geq 0$, reflects the sensitivity to reciprocity utility. The higher ρ_e, the larger the importance of reciprocity utility relative to material utility. If $\rho_e = 0$, an expert seller only considers his own material payoff, and we are trivially back to the standard prediction: the expert seller always undertreats or overcharges the consumer, which leads to market breakdown. If $\rho_e > 0$, reciprocity utility becomes relevant.

Second, $\phi_c(a_c)$ quantifies the extent to which the expert seller perceives a_c as a kind action. As discussed in Falk and Fischbacher (2006), this is the case if a_c increases the expected material payoff of the expert seller $\pi_e(a_e, a_c)$ relative to a reference payoff $\overline{\pi}_e$.[15] In our setting, a natural reference for expert sellers to evaluate the kindness of a_c is the equitable payoff which occurs when the consumer opts out of the market ($\overline{\pi}_e = \overline{\pi}_c = 1.6$).[16] In turn, any action by the consumer allowing the expert seller to earn more than the outside option is perceived as kind. For example, if a consumer decides to interact with the expert seller, and the expert seller applies payoff-maximizing strategy (q_l, p_h), the corresponding kindness term is given by: $\phi_c(a_c = \textit{interaction}) = p_h - c_l - o$. Since a self-interested expert seller is expected to post $p_h \geq 4$, the kindness term is positive. Instead, if a consumer decides not to interact, the kindness term is zero, and reciprocity utility becomes irrelevant.

In treatments GE and GEC, conditionally on the decision to interact, the consumer further decides whether to gift the expert seller. In GE, an unconditional gift increases the maximum expected payoff of the expert seller by $x = 1$. Under the assumption that the expert seller applies payoff-maximizing strategy (q_l, p_h), the kindness term is given by: $\phi_c(a_c = \textit{gift in GE}) = p_h - c_l + x - o > 0$. By contrast, in the GEC treatment the gift is transferred only when sufficient quality is provided, so that: $\phi_c(a_c = \textit{gift in GEC}) = p_h - c_l + (1 - h) \cdot x - o > 0$. Sending a gift is therefore unambiguously perceived as kind in both GE and GEC, although the kindness term

---

[15] In Falk and Fischbacher (2006), an action by the consumer could be unintentional, which affects perceived kindness. In our experiment, all actions are treated as intentional, so that that we do not discuss the intention factor and focus on the material consequences of action a_c by the consumer.

[16] Many papers investigating reciprocity assume that the equitable payoff serves as a reference to assess the kindness of one's action (see e.g. Fehr and Schmidt, 1999; Charness and Rabin, 2002; Cox et al., 2007; Charness and Shmidov, 2014). Apart from the outside option, equitable payoffs are for example generated if an expert seller is honest and always supplies the adequate service, posts the price vector $(p_h, p_l) = (4, 8)$ and the consumer chooses to interact. In contrast to this special case, the outside option serves as a more natural reference point in our context. The implications of our model, however, do not depend on the choice of the outside option as a reference payoff because the decision to interact always allows the expert seller to choose actions to increase their payoff over that of the consumer.

is lower in GEC. This is consistent with experimental evidence on backfiring sanctions or minimum performance requirements in a broader principal-agent context (Fehr and Rockenbach, 2003; Fehr and List, 2004; Falk and Kosfeld, 2006).[17]

The third component of the model, the reciprocation term $\sigma_e(a_e)$, measures how much the expert seller increases the payoff of the consumer in response to a kind action. Relative to profit maximizing strategy (q_l, p_h), for which $\sigma_e = 0$, the expert seller can engage in three types of behavior to increase the consumer's payoff. First, if the consumer needs q_h, the expert seller can abstain from undertreatment and provide q_h. This increases the consumer's payoff by $\sigma_e(a_e = \textit{no undertreatment}) = v$. Second, for a consumer who needs and receives q_h, the expert seller may charge p_l instead of p_h. This implies an increase in the consumer's payoff by $\sigma_e(a_e = \textit{undercharging}) = v + (p_h - p_l)$. In our context, undercharging is akin to a discount and is the strongest form of reciprocal behavior by the expert seller. Lastly, if the consumer needs q_l, the expert seller can abstain from overcharging by applying p_l rather than p_h. The reciprocation term is given by $\sigma_e(a_e = \textit{no overcharging}\ ) = p_h - p_l$.

Based on this framework, we now formulate the implications as a set of hypotheses about the effect of a gift on reciprocal behavior of the expert seller. The first immediately follows from the presence of reciprocal expert sellers.

Hypothesis 1a. *Sending the gift increases the fraction of expert sellers who behave in a consumer-friendly manner.*

The second hypothesis is implied by the fact that an unconditional gift has a higher impact on the expected payoff of the expert seller relative to a conditional gift, so that the kindness term is larger for a gift in the GE treatment as compared to GEC.

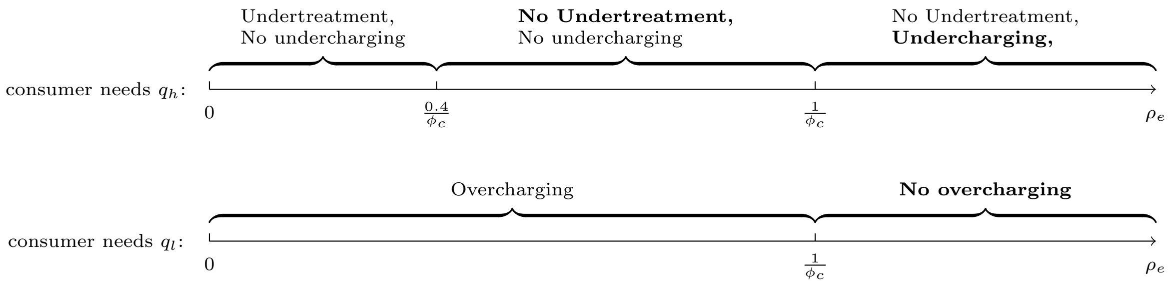
Hypothesis 1b. *An unconditional gift in the GE treatment induces a larger reciprocal response by expert sellers relative to a conditional gift in the GEC treatment.*

Next, how the expert seller reciprocates to a kind action by the consumer depends on the sensitivity parameter ρ_e, which is likely heterogeneous in the population (see e.g. Tang, 2020). In Figure 3, we depict how the different reciprocal actions of the expert seller

---

[17] Several mechanisms behind the backfiring effect of imposing conditions on agents have been discussed, *inter alia* signaling lower trust or communicating lower expectations. Without excluding these channels, we model that conditional gifts are perceived less kind due to a lower impact on the payoff of the expert seller.

Figure 3: Reciprocal response by expert sellers as a function of ρ_e



Notes: The reciprocity parameter ρ_e measures the expert seller's sensitivity to reciprocity utility. The kindness term ϕ_c measures the kindness of the consumer's action as perceived by the expert seller. The gift increases ϕ_c and shifts the respective thresholds leftwards, therefore reducing the likelihood of undertreatment and overcharging and increasing that of undercharging.

(abstaining from undertreatment, undercharging, and abstaining from overcharging) depend on ρ_e, for a given action a_c and associated kindness term ϕ_c. Formally, when a consumer needs q_h, the expert seller abstains from undertreatment whenever $p_h - c_h + \rho_e \cdot \phi_c \cdot v > p_h - c_l \Leftrightarrow \rho_e > \frac{0.4}{\phi_c}$, and further undercharges if $p_l - c_h + \rho_e \cdot \phi_c \cdot (v + (p_h - p_l)) > p_h - c_h + \rho_e \cdot \phi_c \cdot v \Leftrightarrow \rho_e > \frac{1}{\phi_c}$. Similarly, when a consumer needs q_l, the expert seller abstains from overcharging if $p_l - c_l + \rho_e \cdot \phi_c \cdot (p_h - p_l) > p_h - c_l \Leftrightarrow \rho_e > \frac{1}{\phi_c}$.

As an implication of Figure 3, we formulate a hypothesis on the frequency of specific actions by expert sellers.

Hypothesis 2. *In the presence of reciprocating expert sellers, kind actions by consumers have the largest impact on the rate of undertreatment, followed by overcharging and undercharging.*

This hypothesis is in line with experimental evidence showing that agents reciprocate more if their action has a higher relevance for the principal's outcome (Gneezy, 2005; Hennig-Schmidt et al., 2010; Montinari et al., 2016; Englmaier and Leider, 2020). Moreover, undertreatment can be observed by consumers (the problem is not solved), and this can also affect the extent of reciprocity (Güth et al., 1996; Andreoni and Bernheim, 2009; Hoppe and Schmitz, 2018). Results from Dulleck et al. (2011) show that undertreatment occurs less often than overcharging. We note, however, that the possibility of undercharging is not discussed in previous studies on credence goods.

Turning to the consumers, if they anticipate that the expert seller will reciprocate a kind action a_c, this can be expected to motivate both market participation and gifting.

Hypothesis 3a. *The possibility to gift increases interactions in GE and GEC treatments as compared to BASE.*

Hypothesis 3b. *In GE and GEC treatments, a positive fraction of consumers gifts the expert seller.*

In addition, if consumers expect a higher reciprocal response when they transfer an unconditional gift (GE treatment), they will interact and gift more in GE. However, since the conditional gift in the GEC treatment may not be transferred, the expected cost for consumers is lower. In turn, the difference in gifting between GE and GEC treatments is indeterminate.

Lastly, if the GE and GEC treatments lead to more consumer-friendly behavior and more interactions (Hypotheses 1a, 3a, and 3b), the payoffs of consumers and expert sellers would be on average larger in GE and GEC treatments relative to BASE. Defining market efficiency as the sum of profits of consumers and expert sellers (divided by the sum of maximum potential payoff, both normalized by the outside option, see Dulleck et al., 2011), the possibility to gift can be expected to mitigate market inefficiencies associated with asymmetric information.

Hypothesis 4. *Conditional on Hypotheses 1a, 3a, and 3b, profits of consumers and expert sellers are higher in GE and GEC relative to BASE. In turn, the possibility to gift increases market efficiency relative to BASE.*

We close this section by noting that overtreatment is a dominated strategy for expert sellers regardless of their preferences for reciprocity ρ_e. In particular, expert sellers who overtreat supply q_h when the consumer needs q_l, and thereby reduce their own payoff by $c_h - c_l = 4$. Moreover, overtreatment does not increase the payoff of the consumer, so that it does not represent a reciprocal action. Instead, Kerschbamer et al. (2017) suggests that overtreatment is consistent with inequality aversion by expert sellers (see also Beck et al., 2014, for further discussion of overtreatment).

4 Results

Experimental results aggregated at the level of matching groups are reported in Table 1. Following Dulleck et al. (2011), we complement these with a set of random effects panel regressions specified at the level of individual subjects, which notably allows us

Table 1: Overview of experimental results across treatments

	BASE	Gift exchange (GE)			Conditional gift exchange (GEC)		
		Total	Gift	No gift	Total	Gift	No gift
Undertreatment ^{1,2}	0.47	0.49	0.36	0.54	0.48	0.40 ^{a}	0.54 ^{a}
Overcharging ^{1,3}	0.63	0.48	0.43	0.51	0.57	0.59	0.55
Undercharging ^{1,4}	0.09	0.09	0.22	0.04	0.12	0.15	0.09
Overtreatment ^{1,5}	0.19	0.25	0.41 ^{a}	0.17 ^{a}	0.24	0.25	0.23
Interaction ^{1}	0.56	0.60	1.00	1.00	0.65	1.00	1.00
Gift ^{1}	0.00	0.21	1.00	0.00	0.29	1.00	0.00
Profit expert sellers ^{6}	2.73	2.95	3.05	2.93	2.83	2.84	2.93
Profit consumers ^{6}	1.12	0.78	0.57	0.88	1.14	1.22	1.03
Efficiency ^{7}	0.23	0.19	0.15	0.22	0.28	0.31	0.27
p_l posted ^{6}	4.40	4.84	4.23 ^{a}	4.65 ^{a}	4.39	3.89 ^{a}	4.38 ^{a}
p_h posted ^{6}	8.02	7.98	7.61	7.67	7.56	6.84 ^{a}	7.42 ^{a}
Participants in markets	56	56	56	56	56	56	56

Notes: 1) relative frequency; 2) consumer needs q_h, but expert seller provides q_l; 3) consumer needs q_l, receives q_l but is charged p_h and $p_h > p_l$; 4) consumer needs q_h, receives q_h but is charged p_l and $p_h > p_l$; 5) consumer needs q_l and receives q_h; 6) in experimental currency units (1 point= CHF 0.5) and 7) calculated as: *(actual average profit-outside option)/(maximum average profit - outside option)*. For columns BASE, GE (Total) and GEC (Total), Mann-Whitney U-tests for pairwise differences between treatments show no statistical significance at p<0.05 (matching groups of eight subjects are treated as independent observations). For columns GE (Gift), GE (No gift), GEC (Gift) and GEC (No gift), two-tailed Wilcoxon signed-rank tests for pairwise differences with p<0.05 are denoted with superscript *a* (matching groups of eight subjects are treated as independent observations).

to control for dynamic effects (e.g. learning).[18] Formally, the specification we consider can be written as:

$$\mathrm{Y}_{it} = \beta_0 + \beta_1\,\mathrm{GE}_i + \beta_2\,\mathrm{GEC}_i + \beta_3\,\mathrm{GE}_i \text{ x } \mathrm{Gift}_{it} + \beta_4\,\mathrm{GEC}_i \text{ x } \mathrm{Gift}_{it} + \beta_5'\mathrm{X}_{it} + a_i + u_{it}, \quad (2)$$

where Y_{it} is the dependent variable for subject i in period t, GE_i and GEC_i are binary treatment indicators, and Gift_{it} is an indicator variable equal to one if the consumer transferred a gift in period t, zero otherwise. The vector of control variables, denoted X_{it}, includes time period fixed effects and we also check for the sensitivity of our results to the inclusion of prices posted by the expert sellers. Lastly a_i are random effects and u_{it} is a random error term. Throughout we report standard errors clustered at the session level.

In the following, we identify six key results and sequentially discuss evidence pertaining to undertreatment, undercharging, overcharging, market interactions, gifting behavior, as well as profits and market efficiency.

---

[18] Note that some of the outcome variables we consider are binary, and for ease of interpretation we apply a linear probability model. Results are consistent for non-linear models (e.g. probit).

Table 2: Random effects regressions for undertreatment

	Outcome: undertreatment=1			
	(1)	(2)	(3)	(4)
Main effects:				
GE treatment	−0.038	−0.030	0.029	0.039
	(0.059)	(0.053)	(0.049)	(0.043)
GEC treatment	−0.018	−0.028	0.027	0.032
	(0.108)	(0.100)	(0.110)	(0.102)
GE x Gift	–	–	−0.180*	−0.183*
			(0.102)	(0.096)
GEC x Gift	–	–	−0.102**	−0.147***
			(0.048)	(0.055)
Controls:				
Period FEs	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes
# Observations	393	393	393	393

Notes: Random effects panel regressions for expert sellers reported. The outcome variable is a binary indicator equal to one if an expert seller undertreats the consumer, zero otherwise. Column (1) identifies the average treatment effect for GE and GEC treatments. In column (2), we extend the first column controlling for prices. In column (3), we introduce interaction terms for treatments with a variable Gift_{it} equal to one if the consumer transferred a gift in period t, zero otherwise. Column (4) extends column (3) by controlling for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

Result 1. A gift in both GE and GEC treatments reduces the likelihood of undertreatment.

Undertreatment. Table 1, row 1, reports aggregate undertreatment measured as the share of interactions in which expert sellers provide q_l to consumers in need of q_h.[19] In all three treatments, the undertreatment rate is approximately 50%. When expert sellers receive an unconditional gift (GE treatment) undertreatment declines to 36%, which is 18 percentage points lower than when no gift is transferred. When a conditional gift is offered (GEC treatment), undertreatment occurs in 40% of interactions, which is 14 percentage points lower than in the absence of a gift.

Table 2 reports regression results for equation (2), where the outcome variable is equal to 1 if the expert seller undertreats the consumer, zero otherwise. Columns (1) and (2) confirm that the likelihood of undertreatment is not significantly different across

---

[19] As in Dulleck et al. (2011), we restrict the analysis to observations for which undertreatment could potentially occur, that is when consumers interact and need q_h.

Table 3: Random effects regressions for overcharging

	Outcome: overcharging=1			
	(1)	(2)	(3)	(4)
Main effects:				
GE treatment	−0.112	−0.100	−0.092	−0.071
	(0.138)	(0.143)	(0.135)	(0.137)
GEC treatment	−0.063	−0.055	−0.069	−0.064
	(0.089)	(0.093)	(0.081)	(0.085)
GE x Gift	–	–	−0.063***	−0.089***
			(0.018)	(0.019)
GEC x Gift	–	–	0.013	0.017
			(0.062)	(0.070)
Controls:				
Period FEs	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes
# Observations	393	393	393	393

Notes: Random effects panel regressions for expert sellers reported. The outcome variable is a binary indicator equal to one if an expert seller overcharges the consumer, zero otherwise. Column (1) identifies the average treatment effect for GE and GEC treatments. In column (2), we extend the first column controlling for prices. In column (3), we introduce interaction terms for treatments with a variable Gift_{it} equal to one if the consumer transferred a gift in period t, zero otherwise. Column (4) extends column (3) by controlling for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

treatments. Column (3) further shows that gifting significantly reduces undertreatment, albeit at a slightly higher rate in GE. This is in line with Hypotheses 1a and 1b. Controlling for prices (column 4) does not affect our conclusions.

Result 2. An unconditional gift by consumers in GE decreases the likelihood of overcharging and increases the likelihood of undercharging.

Overcharging. The rate of overcharging is defined as the share of interactions in which consumers need and receive q_l, but are charged p_h.[20] The second row of Table 1 shows that the overcharging rate is 63% in BASE, 48% in GE, and 57% in GEC, although the differences are not statistically significantly different from zero. Estimates based on equation 2 for a binary outcome variable equal to one if overcharging occurs confirm this (Table 3, columns 1 and 2).

---

[20] We only consider observations for which overcharging can occur, that is when the consumer interacts, needs q_l and the price vector implies $p_h > p_l$ (see Dulleck et al., 2011). We discard 25 observations with $p_h = p_l$.

More interestingly, our results provide some evidence that an unconditional gift (GE treatment) reduces the likelihood that consumers are charged for a service they did not receive, as the aggregate overcharging rate declines by eight percentage points. In regression results both interaction terms are statistically significantly different from zero and suggest that the decline in the probability of overcharging induced by the gift is between six and nine percentage points (Table 3, columns 3 and 4). By contrast, results for the GEC treatment suggest that the impact of a conditional gift is close to zero. These results are consistent with Hypotheses 1b and 2.

Undercharging. The rate of undercharging, defined as the proportion of interactions in which consumers need and receive q_h, but are charged the price of the low-quality service p_l, is around 10% in all three treatments (Table 1, row 3).[21] In the GE treatment, a gift by consumers increases undercharging by about 18 percentage points. In the GEC treatment, sending a conditional gift only increases undercharging by six percentage points.

Table 4 shows corresponding regression results for a binary outcome variable equal to one if undercharging occurs, zero otherwise. Results confirm that unconditional gifting (GE treatment) significantly increases undercharging (columns 3 and 4). Specifically, when an unconditional gift is transferred, the likelihood of undercharging increases by around 17 percentage points. By contrast, the coefficient associated with a gift in the GEC treatment is small and statistically insignificant. This is again in line with Hypotheses 1b and 2.

Result 3. The share of consumers who choose to interact in the credence goods market is similar in all three treatments.

Interactions. The fifth row of Table 1 shows that consumers decide to interact in 56% of the cases in BASE, 60% in GE, and 65% GEC. While the possibility to gift is therefore associated with an increase in interactions, both non-parametric tests and regression results reported in Table 5, columns (1) and (2), suggest that the differences across treatments are not statistically significant. This result goes against Hypothesis 3a.

Result 4. The proportion of consumers who send the gift is similar in GE and GEC treatments.

---

[21] We again consider observations for which undercharging can occur, that is when the consumer interacts, needs q_h and the price vector implies $p_h > p_l$ (Dulleck et al., 2011). We drop 12 observations with $p_h = p_l$.

Table 4: Random effects regressions undercharging

	Outcome: undercharging=1			
	(1)	(2)	(3)	(4)
Main effects:				
GE treatment	0.005	−0.001	−0.059**	−0.064**
	(0.027)	(0.026)	(0.024)	(0.028)
GEC treatment	0.039	0.025	0.028	0.018
	(0.030)	(0.035)	(0.037)	(0.042)
GE x Gift	–	–	0.172***	0.168***
			(0.062)	(0.064)
GEC x Gift	–	–	0.024	0.018
			(0.025)	(0.034)
Controls:				
Period FEs	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes
# Observations	381	381	381	381

Notes: Random effects panel regressions for expert sellers reported. The outcome variable is a binary indicator equal to one if an expert seller undercharges the consumer, zero otherwise. Column (1) identifies the average treatment effect for GE and GEC treatments. In column (2), we extend the first column controlling for prices. In column (3), we introduce interaction terms for treatments with a variable Gift_{it} equal to one if the consumer transferred a gift in period t, zero otherwise. Column (4) extends column (3) by controlling for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

Gifting. The frequency of gifting is reported in the sixth row of Table 1, and indicates that 21% of consumers send a gift in the GE treatment, 29% in the GEC treatment. Conditional on interaction, this corresponds to 35% and 45% for GE and GEC respectively. A positive propensity to gift is in line with Hypothesis 3b. Regression results reported in Table 5, columns (3) and (4), suggest that differences in the probability to send the gift are not statistically different in GE and GEC (Wald test for the equality of GE and GEC estimates in column 3: $p = 0.178$).

Result 5. The possibility to gift in GE and GEC treatments does not significantly increase market efficiency relative to BASE.

Result 6. For consumers who need q_h, sending a gift significantly reduces market inefficiencies.

Profits and Market efficiency. The seventh and eighth rows in Table 1 show that profits

Table 5: Random effects regressions for interaction and gifting

	Outcome: interaction=1		Outcome: gift=1	
	(1)	(2)	(3)	(4)
Main effects:				
GE treatment	0.042	0.049	0.214***	0.219***
	(0.048)	(0.072)	(0.049)	(0.054)
GEC treatment	0.087	0.048	0.286***	0.269***
	(0.056)	(0.084)	(0.019)	(0.023)
Controls:				
Period FEs	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes
# Observations	1,344	1,344	1,344	1,344

Notes: Random effects panel regressions for consumers reported. In columns (1) and (2) the outcome variable is a binary indicator equal to one if the consumer interacts, zero otherwise. In columns (3) and (4) the outcome variable is a binary indicator equal to one if the consumer gifts the expert seller, zero otherwise. Columns (2) and (4) include prices as control variables. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

of expert sellers are on average very similar across treatments. The profits of expert sellers are also not significantly affected when consumers send a gift. Similarly, while the profits for consumers tend to be lower than those earned by expert sellers, there are no significant differences across treatments. In line with this, market efficiency reported in row eight shows no significant differences across columns. This result contradicts Hypothesis 4.

To provide further evidence on how gifting and reciprocity affect profits and market efficiency, we focus on periods in which the consumer participates in the market and distinguish cases in which either q_h or q_l is needed.[22] Starting with profits for expert sellers, panel regression results reported in Table 6 confirm that treatment effects for GE and GEC are not statistically significantly different from zero (columns 1 and 2 for q_h, 5 and 6 for q_l).

More importantly, both conditional and unconditional gifts by a consumer who needs

---

[22] As should be clear from above, observations for consumers who need q_h are those for which undertreatment and undercharging are possible (respectively Table 2 and Table 4), except for cases with $p_h = p_l$. Similarly, results for q_h correspond to the sample where overcharging is possible (Table 3) except for observations with $p_h = p_l$.

Table 6: Random effects regressions for expert sellers' profits

	Consumer needs q_h				Consumer needs q_l			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main effects:								
GE treatment	0.193	0.198	0.314	0.282^{*}	0.340	0.068	0.253	-0.014
	(0.188)	(0.154)	(0.201)	(0.152)	(0.542)	(0.487)	(0.572)	(0.477)
GEC treatment	-0.180	0.087	0.024	0.072	-0.086	0.179	-0.603^{*}	-0.466^{*}
	(0.217)	(0.387)	(0.252)	(0.435)	(0.301)	(0.311)	(0.312)	(0.269)
GE x Gift	–	–	-0.320	-0.221	–	–	0.273^{**}	0.212^{*}
			(0.314)	(0.376)			(0.128)	(0.128)
GEC x Gift	–	–	-0.473^{*}	0.038	–	–	1.130^{***}	1.419^{***}
			(0.286)	(0.324)			(0.099)	(0.280)
Controls:								
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes	No	Yes	No	Yes
# Observations	393	393	393	393	418	418	418	418

Notes: Random effects panel regressions for expert sellers reported. The outcome variable is the profits of the expert sellers conditional on consumer needing q_h (columns 1-4) or q_l (columns 5-8). Columns (1) and (5) identify the average treatment effect for GE and GEC treatments. Columns (2) and (6) additionally control for prices. In columns (3) and (7), we extend columns (1) and (5) respectively adding interaction terms for treatments with a variable Gift$_{it}$ equal to one if the consumer transferred a gift in period t, zero otherwise. In columns (4) and (8) we additionally control for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

q_h do not significantly increase expert sellers' average profits (columns 3 and 4), even though expert sellers receive one additional currency unit (Wald test that interaction terms are equal to one: $p < 0.01$). This is in line with the observation that expert sellers abstain from undertreating consumers and also undercharge, thereby giving up some of their own profits to reciprocate the gift. By contrast, when the consumer needs q_l (Table 6, columns 7 and 8), both types of gift have a positive impact on the profits of expert sellers. We note, however, that an unconditional gift has a smaller impact on expert seller's profits relative to a conditional gift (Wald test for the equality of coefficients is rejected with $p < 0.01$). This corresponds to the observation that overcharging declines only for an unconditional gift.

Turning to consumers' profits, Table 7 columns 1-2 and 5-6 show that GE and GEC treatments tend to be detrimental to consumers on average. However, columns 3 and 4 show that, if the consumer needs q_h, sending either type of gifts has a positive impact on consumer's profits. This suggests that the direct cost of the gift for consumers is more than compensated by an increase in consumer-friendly behavior by expert sellers. When the consumer needs q_l, however, columns 7 and 8 show that sending the gift has a significant negative impact on profits. In other words, the decline in overcharging is

Table 7: Random effects regressions for consumers' profits

	Consumer needs q_h				Consumer needs q_l			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main effects:								
GE treatment	-0.568^{*}	-0.691^{**}	-1.042^{***}	-1.100^{***}	-0.443	-0.140	-0.083	0.139
	(0.339)	(0.314)	(0.393)	(0.380)	(0.368)	(0.263)	(0.473)	(0.308)
GEC treatment	-0.018	-0.327	-0.558	-0.637	-0.106	-0.475^{**}	0.319	0.152
	(0.820)	(0.916)	(0.821)	(0.983)	(0.286)	(0.185)	(0.330)	(0.171)
GE x Gift	–	–	1.240^{*}	1.088	–	–	-1.064^{***}	-0.751^{***}
			(0.674)	(0.688)			(0.302)	(0.169)
GEC x Gift	–	–	1.274^{***}	0.778	–	–	-0.964^{***}	-1.407^{***}
			(0.432)	(0.490)			(0.150)	(0.157)
Controls:								
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes	No	Yes	No	Yes
# Observations	393	393	393	393	418	418	418	418

Notes: Random effects panel regressions for consumers reported. The outcome variable is the profits of the consumer conditional on needing q_h (columns 1-4) or q_l (columns 5-8). Columns (1) and (5) identify the average treatment effect for GE and GEC treatments. Columns (2) and (6) additionally control for prices. In columns (3) and (7), we extend columns (1) and (5) respectively adding interaction terms for treatments with a variable Gift_{it} equal to one if the consumer transferred a gift in period t, zero otherwise. In columns (4) and (8) we additionally control for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

not sufficient to compensate the cost of the gift.

Lastly, regression results for market efficiency are reported in Table 8, confirming that the treatment effect of GE and GEC are small and not statistically significant (columns 1-2 and 5-6). When consumers need q_h, however, sending both types of gift significantly increases market efficiency (columns 3 and 4). Since undercharging is a transfer and therefore does not directly affect market efficiency, this effect can instead be attributed to the observed decline in undertreatment.

Table 8: Random effects regressions for market efficiency

	Consumer needs q_h				Consumer needs q_l			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main effects:								
GE treatment	-0.241	-0.273	-0.703^{**}	-0.733^{**}	-0.052	-0.067	0.018	-0.004
	(0.342)	(0.325)	(0.329)	(0.331)	(0.068)	(0.071)	(0.048)	(0.060)
GEC treatment	-0.138	-0.072	-0.426	-0.541	-0.060	-0.068^{*}	-0.065^{*}	-0.077^{**}
	(0.791)	(0.732)	(0.749)	(0.696)	(0.051)	(0.041)	(0.039)	(0.035)
GE x Gift	–	–	1.221^{**}	1.223^{**}	–	–	-0.209^{***}	-0.183^{***}
			(0.490)	(0.486)			(0.049)	(0.046)
GEC x Gift	–	–	0.683^{***}	1.170^{***}	–	–	0.011	0.022
			(0.182)	(0.268)			(0.049)	(0.045)
Controls:								
Period FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes	No	Yes	No	Yes
# Observations	393	393	393	393	418	418	418	418

Notes: Random effects panel regressions for consumers reported. The outcome variable is total market efficiency conditional on consumers needing q_h (columns 1-4) or q_l (columns 5-8). Efficiency is calculated as the sum of profits of interacting expert sellers and consumers, divided by the maximum average profit (=2 in condition h and =4 in condition l), both normalized by the outside option. Columns (1) and (5) identify the average treatment effect for GE and GEC treatments. Columns (2) and (6) additionally control for prices. In columns (3) and (7), we extend columns (1) and (5) respectively adding interaction terms for treatments with a variable Gift_{it} equal to one if the consumer transferred a gift in period t, zero otherwise. In columns (4) and (8) we additionally control for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

Table 9: Random effects regressions for overtreatment

	Outcome: overtreatment=1			
	(1)	(2)	(3)	(4)
Main effects:				
GE treatment	0.039	0.049	−0.017	−0.004
	(0.094)	(0.097)	(0.076)	(0.085)
GEC treatment	0.070	0.076	0.083	0.088*
	(0.060)	(0.049)	(0.053)	(0.049)
GE x Gift	–	–	0.176***	0.165***
			(0.050)	(0.049)
GEC x Gift	–	–	−0.025	−0.026
			(0.053)	(0.047)
Controls:				
Period FEs	Yes	Yes	Yes	Yes
Prices	No	Yes	No	Yes
# Observations	418	418	418	418

Notes: Random effects panel regressions for expert sellers reported. The outcome variable is a binary indicator equal to one if an expert sellers overtreats the consumer, zero otherwise. Column (1) identifies the average treatment effect for GE and GEC treatments. In column (2), we extend the first column controlling for prices. In column (3), we introduce interaction terms for treatments with a variable Gift$_{it}$ equal to one if the consumer transferred a gift in period t, zero otherwise. In column (4) we extend the third column controlling for prices. All specifications include period fixed effects. Robust standard errors clustered at the session level reported in parentheses. *, ** and *** denote statistical significance at 10%, 5% and 1% respectively.

When the consumer needs q_l, sending an unconditional gift (GE treatment) has a negative and statistically significant impact on market efficiency (Table 8, columns 7 and 8), whereas a conditional gift (GEC treatment) has a small and statistically insignificant impact. This change in market efficiency is not driven by the change in overcharging, which represents a transfer. Instead, we show in Table 9 that the decline in market efficiency can be explained by an increase in overtreatment when an unconditional gift is sent (GE treatment, columns 3 and 4).[23] As discussed previously, overtreatment reduces the payoff of the expert seller, but it does not affect consumers' payoff. This response to a gift is therefore not consistent with the model by Falk and Fischbacher (2006), and suggests the presence of inequality averse expert sellers as discussed by Kerschbamer et al. (2017).

Taken together, these results indicate that sending a gift has a positive impact on

---

[23] The analysis focuses on observations for which overtreatment can occur, that is when the consumer chooses to interact and needs q_l.

consumer's profits and efficiency in markets where consumers face a severe problem and need q_h. By contrast, when q_l is needed, gifting reduces the consumer's profit and can also reduce market efficiency. These effects cancel each other, so that introducing a possibility to gift expert sellers does not affect overall market efficiency.

5 Discussion and conclusion

Using a canonical experimental market for credence goods, this paper has introduced the possibility for consumers to send conditional and unconditional gifts to expert sellers, and quantified implications for the behavior of expert sellers as well as for overall market efficiency. Our results confirm that sending an unconditional gift triggers more consumer-friendly behavior by expert sellers, as the rate of undertreatment and overcharging decline, while undercharging increases. We also find that conditional gifts only reduce undertreatment, which suggests that expert sellers perceive these as less kind. Conditioning a gift, which is akin to minimum performance contracts, can therefore backfire, as it fails to trigger reciprocal actions that cannot be observed by consumers.

While our results provide novel evidence on the importance of reciprocity for credence goods markets, they also suggest that the possibility to gift expert sellers does not significantly increase overall market efficiency. However, we show that the benefit of gifting depends on the severity of the problem faced by the consumer, as market efficiency improves for high-severity consumers. We also note that, in our experiment, interaction rates in all treatments remain relatively low, and consumers need q_h in only 50 percent of the cases, which implies that the scope for gift exchange to significantly increase market efficiency is limited.

In light of this, we emphasize one critical feature of the credence goods market in our study: the severity of the problem faced by consumers is manipulated experimentally. In other words, the consumers we consider have no private information about their own type. In settings where consumers have some information about the severity of the problem they face, offering the expert seller a gift may be beneficial. Investigating the impact of such strategies constitutes an interesting area for future research.

Conclusion

In this thesis, I have investigated the credence component of energy-transforming technologies. In particular, I have provided a theoretical framework showing that energy-transforming technologies have a credence component (Chapter I), empirical evidence related to second-degree moral hazard in markets for subsidized solar photovoltaic systems (Chapter II) as well experimental evidence on how to mitigate these (Chapter III).

The main findings of this thesis provide important insights for climate policies related to the adoption of energy-efficient technologies. In essence, accounting for the credence component of energy technologies and related supply-side inefficiencies is key to promote their adoption in a cost-effective way. Particularly, the demand-side needs to be informed about their needs and technologies as much as possible, trust in certification needs to be established and the work of expert sellers should be verified whenever possible.

In the first chapter, I have studied the credence component of energy-transforming technologies and shown that the credence goods framework can be useful to identify barriers to the adoption of energy-efficient technologies. Combining a basic model of energy efficiency investments with the seminal model for credence goods, I have identified how asymmetric information related to energy-technologies can lead to three types of supply side inefficiencies and low market participation. Difficulties to quantify energy savings and/or energy efficiency lead to the failure of verifiability and liability to improve market outcomes. In sum, the results from Chapter I suggest that insights about inefficiencies inherent to credence goods are useful for researchers and policymakers dealing with the market for energy efficiency.

In the second Chapter, I have investigated whether supply-side inefficiencies identified in Chapter I are relevant for the market for solar photovoltaic systems. In particular, the literature on credence goods suggests that supply-side inefficiencies increase if consumers receive third-party reimbursements. In the context of PV systems, subsidies paid to consumers to foster their adoption could therefore affect self-reported expected electricity output as well as transaction prices.

To study this hypothesis, I analyzed data from a solar subsidy program in California and quantified the relationship of subsidy levels and the design factor as well as the transaction price of PV systems. Employing an instrumental strategy to account for potential self-selection of installers into specific subsidy levels and further controlling for a wide range of potential confounding factors, it is shown that TPO installers increase the self-reported design factor and thereby the total amount of subsidies received for residential systems, unless the reported data is verified during a field inspection. Furthermore, my results suggest that second-degree moral hazard increases the transaction prices of HO systems. Heterogeneous results depending on the consumer sector are further in line with heterogeneous social preferences of installers. For example, when governmental consumers, who are financed by the tax-payer and may therefore be perceived as financially well-endowed, lease a system from a TPO installer, the transaction price per Watt is associated with an increase of 25 to 50 percent for a one dollar increase of the subsidy level.

In the third chapter, I have studied the role of reciprocity and gift exchange in an experimental market for credence goods. In this market, consumers had the possibility to send conditional and unconditional gifts to expert sellers and I quantified how both kinds of gifts affect the behavior of expert sellers as well as market efficiency. The results indicate that unconditional gifts lead to more consumer-friendly behavior, undertreatment and overcharging decline, while some consumers even receive a rebate in the form of undercharging. In line with the literature suggesting that expert sellers perceive conditional gifts as less kind, these only lead to more-consumer friendly behavior which the consumer can observe (i.e. undertreatment).

The possibility to contract over energy savings/output is a crucial feature of energy-efficient technologies (see Chapter I for a discussion and Sorrell, 2007, for a theoretical model). Related implications of payments contingent on sufficient services are studied in Chapter II and III of this thesis and results contribute to the understanding of potentials and risks of energy performance contracting. First, conditioning payments on receiving a predefined level of service-quality (i.e. the GEC condition of the laboratory experiment) has shown to prevent consumers from receiving insufficient services. This confirms that aligning incentives of the supply- and demand-side of energy-efficient technologies may increase consumer welfare. However, the laboratory setup in Chapter III abstracts from important real world features of energy performance contracting which are hard to represent in a laboratory environment. In particular, monetary savings from reduced energy consumption are realized over a long time, and Klinke (2018) provides evidence that economic viability is a significant barrier to the adoption of energy performance contracting, mainly because of the risk associated with future energy savings (e.g. changes in the weather). Therefore, testing the external validity of the laboratory evidence on conditional gifts by studying outcomes of energy-contracting in the field is warranted.

Second, results from Chapter II show that subsidies which are paid on expected electricity output (as opposed to conditioning them on actual electricity output) may cause installers to exaggerate expected electricity output to receive larger subsidies at the expense of the subsidy provider. At the same time results from the literature on the design of subsidy programs suggest that consumers significantly discount subsidy payments if they are paid conditional on the electricity output and thus paid in the future (Burr, 2016; Feger et al., 2017; De Groote and Verboven, 2019). A promising alley for future research endeavors is thus to compare attractiveness and cost-effectiveness of unconditional upfront subsidies and conditional future subsidies to inform the design and welfare effects of subsidy programs.

Apart from the conditions and timing of subsidy payments, this thesis further provides an extensive analysis on how third-party reimbursements such as subsidies may generally affect efficiency in markets for credence goods. In Chapter I, it is argued that evidence from credence goods markets suggests that subsidizing energy-efficient technologies may increase supply-side inefficiencies. It is concluded that empirical evidence on how prices for energy-efficient technologies respond to subsidy policies is missing. In Chapter II, I contribute to this research gap and confirm the conjecture that subsidies may increase supply-side inefficiencies and in particular increase transaction prices paid by consumers. This suggests that the welfare effects of subsidy programs could increase if i) consumers signal to the installer that they have some basic knowledge on the work steps needed to set up a solar photovoltaic system, ii) consumers get a second opinion from a different installer, iii) subsidy programs do not give installers control over the total subsidies received and iv) the work of installers is verified after installation.

Besides the use of subsidies, peer effects on household energy behavior have been identified as an important driver for the adoption of energy-efficient technologies, in particular PV systems (see Wolske et al., 2020, for a recent review). For example, within the peer group of a neighbourhood, the likelihood of additional installations of PV systems increases in the number of current installations. Following Wolske et al. (2020), one underlying mechanism is that early adopters who share their positive experiences may increase the trust in technology and installers within their neighbourhood. Importantly, the source of information and also the information itself need to be trustworthy to create peer effects (Palm, 2017). However, in the absence of verifiability, it is difficult for early adopters to assess the quality and pricing of the service provided, implying that peer effects might be driven by subjective assessments and herding effects.[24] To prevent this, early adopters should at best share credible information for example on realized energy output of their PV system in order to create positive and beneficial peer effects.

Another specificity of energy-transforming technologies is that they are often sold by retailers, which act as intermediaries between manufacturers and consumers. Because the credence component may apply at the manufacturer level, such retailers may not be well informed. For example, when car manufacturers overstate the energy efficiency of vehicles, their uninformed retailers may unintentionally recommend a suboptimal vehicle to their consumers. Moreover, even if intermediaries are well informed about energy characteristics of products, they may choose not to disclose it to a majority of uninterested consumers (Allcott and Sweeney, 2017). As a solution, Milgrom (2008) discusses the role of liability rules for withholding information or even for not providing relevant information which the sales agent should have known, although in practice this may be difficult.

As highlighted in Chapter I, a promising way to increase trust in information about energy efficiency provided by manufacturers is competitor testing where firms test and reveal violations of safety and environmental standards related to products of their competitors to a government regulation authority. This procedure has several promising advantages compared to direct testing by government regulation authorities: First, through their own efforts to comply with product standards, firms have the technology, know-how and staff to efficiently test their competitor's products. For the same reason, producers have a better understanding which technology parts are potentially prone to noncompliance. Furthermore, competitor-testing internalizes testing costs because competing firms have direct commercial incentives to uncover noncompliance of their competitors as this could lead to a reduction of the latter's market power. In contrast, testing by government regulation authorities must be publicly funded. Despite these appealing advantages of competitor testing, empirical evidence on its implications for

---

[24] Related, the recent rise of consumer platforms where consumers can provide uni-or bilateral feedback can increase trust in technology and installers (Tadelis, 2016). In a market for credence goods however, these feedback schemes are again likely to be driven by subjective as opposed to verified and therefore credible assessments.

credence goods markets is missing.

The credence component of energy-efficient technologies has attracted interest of economic research only recently. To answer the question whether possible, costly remedies are welfare increasing more studies are needed to quantify the effect sizes and welfare effects of supply-side inefficiencies originating from the credence component of energy-efficient technologies. Chapter II provides a first explorative quantification of the cost of the welfare consequences related to second-degree moral hazard. While for example an increase of the design factor of HO systems by 0.5 percentage points followed from a one dollar increase of the subsidy level may seem small, mandatory field inspections for every system would have decreased total average subsidy payments by 250 USD per solar photovoltaic system. Furthermore, many of the policy recommendations come at a low cost or even for free. From the consumer's perspective, gathering and signaling information on the adequate technologies to expert sellers or giving them little gifts can help to increase consumer welfare. From the regulator's perspective, giving firms incentives to test the quality of their competitor's technologies may ultimately make the financing of GRAs redundant.

I want to close by emphasizing how the results from this thesis contribute to the wider literature on credence goods. First of all, the experimental results from the third chapter are applicable to the general framework for credence goods and can thus inform other typically quoted markets for credence goods such as cab rides. In such a context, an unconditional tip would be akin to committing to tip when entering the cab and a conditional gift would be akin to committing to tip conditional upon arriving at the destination within a certain time. The experimental results suggest that both types of gifts can circumvent the tedious task of identifying a nice cab driver (Kerschbamer et al., 2017). It would further be interesting to see whether the stronger reciprocal response to the unconditional gift in our laboratory setup is confirmed in a field experiment. In the future, laboratory experiments on gift exchange in markets for credence goods should also study the effect of giving or not giving small gifts in contexts where the gift is either expected (e.g. as part of a cultural norm) or when they come as a surprise.

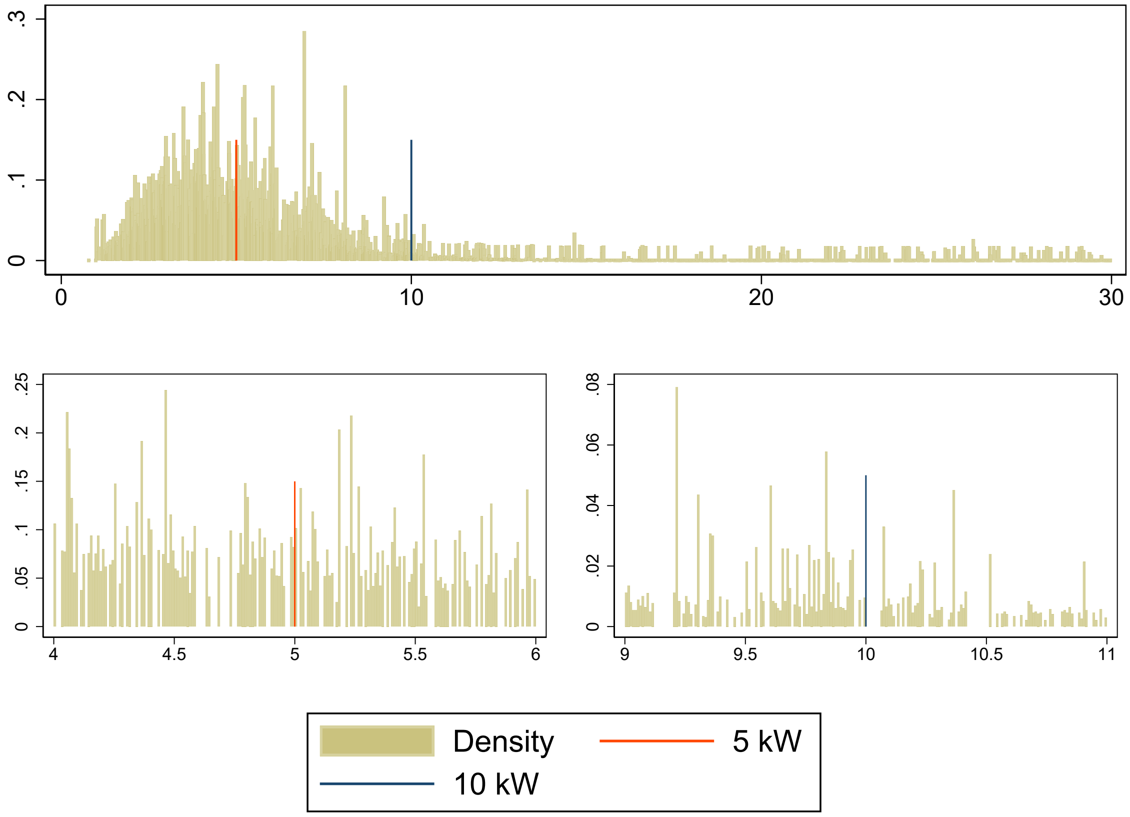
From a more conceptual perspective, this thesis demonstrates that the literature on energy-efficient technologies has come up with several remedies which can be applied to other markets for credence goods. One example is the idea of ex-post verification of an expert seller's service such as the mandatory field inspections of PV systems discussed in Chapter II. If dentists for example knew that health insurances would randomly inspect some of their services and corresponding prices by reviewing the medical records and actual condition of the respective patient, the dentists might provide more adequate and cheaper services in the first place. Linking remedies to specific inefficiencies in markets for energy-efficient technologies to the credence goods framework may therefore expand the impact of research in energy and environmental economics.

Appendix

The Appendix contains supplementary information for Chapters II, and III.

A Distribution of system size (Chapter II)

Figure A1: Size distribution of upfront systems (Chapter II)



Notes: Distribution of system size of upfront systems. The upper panel shows all upfront systems up to 30 kW. The lower left panel shows the distribution of the subset of system sized four to six kW. The lower right panel shows the distribution of the subset of system sized nine to eleven kW. The width of bins set to 0.01 kW.

Figure A1 shows the distribution of system size of upfront systems. As discussed in section 3.1, there is no evidence of bunching around a five or ten kW threshold, suggesting that strategic considerations do not play a role when choosing the system size.

Consumers installing a system sized between ten and 30 kW could choose to receive a different kind of subsidy which is paid conditional on actual electricity output rather then expected output. If, for example, installers would want to maximize the upfront amount of subsidies received one would observe bunching of upfront systems with a size just below the threshold of 30 kW. We do not observe evidence for bunching around this threshold, suggesting that strategic self- selection into either subsidy type does not bias the results.

B First stage-regression results (Chapter II)

Table B1: First stage regression results for Tables 3 and 4

	TPO (Table 3)			HO (Table 4)		
	(1)	(2)	(3)	(4)	(5)	(6)
Z_i	0.883^{***} (0.003)	0.884^{***} (0.003)	0.640^{***} (0.027)	0.849^{***} (0.004)	0.850^{***} (0.004)	0.589^{***} (0.008)
Field inspection (FI)						
FI = 1 × Z_i		0.989^{***} (0.007)			0.965^{***} (0.004)	
Sector						
Government × Z_i			0.950^{***} (0.099)			0.713^{***} (0.022)
Non-Profit × Z_i			0.702^{***} (0.081)			0.755^{***} (0.013)
Residential × Z_i			0.973^{***} (0.006)			0.987^{***} (0.002)
N	67,230	67,230	67,230	69,113	69,113	69,113

Notes: The outcome variable is the actually received subsidy level. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

Table B2: First stage regression results for Tables 5 and 6

	TPO (Table 5)			HO (Table 6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Z_i	0.883^{***}	0.883^{***}	0.641^{***}	0.849^{***}	0.850^{***}	0.589^{***}
	(0.003)	(0.003)	(0.027)	(0.004)	(0.004)	(0.008)
Field inspection (FI)						
FI = 1 × Z_i		0.989^{***}			0.965^{***}	
		(0.007)			(0.004)	
Sector						
Government × Z_i			0.950^{***}			0.713^{***}
			(0.099)			(0.022)
Non-Profit × Z_i			0.702^{***}			0.755^{***}
			(0.081)			(0.013)
Residential × Z_i			0.973^{***}			0.987^{***}
			(0.006)			(0.002)
N	67,230	67,230	67,230	69,113	69,113	69,113

Notes: The outcome variable is the actually received subsidy level. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters, the experience of installers, the relative market power of installers and a measure for local industry concentration. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

C Additional Tables (Chapter II)

Table C1: Design factor of TPO systems (Exclusion window)

	Exclusion window +- two weeks					
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)
Subsidy level	0.398^{*}	0.552^{**}	0.399^{*}	0.552^{**}	0.690	1.208
	(0.156)	(0.172)	(0.156)	(0.172)	(0.781)	(0.709)
Field inspection (FI)						
FI = 1 × Subsidy level			−0.447	−0.221		
			(0.459)	(0.480)		
FI = 1			0.709	0.690		
			(0.449)	(0.447)		
Sector						
Government × Subsidy level					−3.668	−5.581
					(3.165)	(4.416)
Non-Profit × Subsidy level					1.202	2.189
					(1.191)	(1.905)
Residential × Subsidy level					0.400^{*}	0.550^{**}
					(0.156)	(0.172)
Government					3.643	5.272
					(3.062)	(4.044)
Non-Profit					1.207	1.319
					(0.797)	(0.873)
Residential					-1.446^{***}	-1.496^{***}
					(0.437)	(0.437)
N	61,456	61,456	61,456	61,456	61,456	61,456
1st-stage partial F-stat.	-	31145.7; 271.5; 435.5; 1.5e+05	-	62.3; 183.2; 100.3; 2056.5	-	157.6; 171.0; 139.0; 1516.1

Notes: The outcome variable is the design factor of TPO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

Table C2: Design factor of HO systems (Exclusion window)

	Exclusion window +- two weeks						
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	
Subsidy level	−0.016	−0.046	−0.019	−0.049	0.013	−0.017	
	(0.111)	(0.126)	(0.111)	(0.127)	(0.225)	(0.267)	
Field inspection (FI)							
FI = 1 × Subsidy level			0.032	0.019			
			(0.170)	(0.183)			
FI = 1			−0.414*	−0.427**			
			(0.162)	(0.165)			
Sector							
Government × Subsidy level					−0.012	−0.388	
					(0.621)	(0.752)	
Non-Profit × Subsidy level					−0.442	−0.428	
					(0.319)	(0.355)	
Residential × Subsidy level					−0.018	−0.042	
					(0.111)	(0.126)	
Government					−0.805	−0.403	
					(0.816)	(0.981)	
Non-Profit					0.758	0.738	
					(0.389)	(0.415)	
Residential					−0.856***	−0.849***	
					(0.153)	(0.153)	
N	69,113	63,063	63,063	63,063	63,063	63,063	63,063
1st-stage partial F-stat.	-	31145.7; 271.5; 435.5; 1.5e+05	-	62.3; 183.2; 100.3; 2056.5	-	157.6; 171.0; 139.0; 1516.1	

Notes: The outcome variable is the design factor of HO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

Table C3: Transaction price per Watt of TPO systems (Exclusion window)

	Exclusion windiw +- two weeks						
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	
Subsidy level	0.049	−0.008	0.049	−0.008	−0.083	−0.223	
	(0.028)	(0.030)	(0.028)	(0.030)	(0.184)	(0.219)	
Field inspection (FI)							
FI = 1 × Subsidy level			0.044	−0.028			
			(0.151)	(0.151)			
FI = 1			−0.088	−0.084			
			(0.111)	(0.111)			
Sector							
Government × Subsidy level					2.877*	3.411***	
					(1.133)	(1.023)	
Non-Profit × Subsidy level					0.589*	0.881	
					(0.300)	(0.460)	
Residential × Subsidy level					0.046	−0.008	
					(0.029)	(0.030)	
Government					−0.578	−1.057	
					(0.802)	(0.722)	
Non-Profit					0.353	0.417	
					(0.222)	(0.237)	
Residential					0.205	0.210	
					(0.128)	(0.129)	
N	69,113	61,456	61,456	61,456	61,456	61,456	61,456
1st-stage partial F-stat.	-	31145.7; 271.5; 435.5; 1.5e+05	-	62.3; 183.2; 100.3; 2056.5	-	157.6; 171.0; 139.0; 1516.1	

Notes: The outcome variable is the transaction price per Watt of TPO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters, the experience of installers, the relative market power of installers and a measure for local industry concentration. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

Table C4: Transaction price per Watt of HO systems (Exclusion window)

	Exclusion window +- two weeks						
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	
Subsidy level	0.270***	0.298***	0.269***	0.297***	0.429***	0.416***	
	(0.037)	(0.047)	(0.037)	(0.047)	(0.097)	(0.105)	
Field inspection (FI)							
FI = 1 × Subsidy level			0.315***	0.333***			
			(0.055)	(0.062)			
FI = 1			−0.077*	−0.070			
			(0.037)	(0.038)			
Sector							
Government × Subsidy level					0.767**	0.671*	
					(0.248)	(0.302)	
Non-Profit × Subsidy level					0.677***	0.699***	
					(0.109)	(0.117)	
Residential × Subsidy level					0.282***	0.295***	
					(0.038)	(0.047)	
Government					0.357	0.463	
					(0.298)	(0.362)	
Non-Profit					−0.555***	−0.576***	
					(0.118)	(0.128)	
Residential					−0.136*	−0.141*	
					(0.056)	(0.055)	
N	69,113	63,063	63,063	63,063	63,063	63,063	63,063
1st-stage partial F-stat.	-	31145.7; 271.5; 435.5; 1.5e+05	-	62.3; 183.2; 100.3; 2056.5	-	157.6; 171.0; 139.0; 1516.1	

Notes: The outcome variable is the transaction price per Watt of HO systems. All specifications include fixed effects for the IOU, installer, month, sector as well as for make and models of modules and inverters. Further, all specifications include controls for the amount of modules and inverters, the experience of installers, the relative market power of installers and a measure for local industry concentration. Robust standard errors clustered at the zip code level are reported in parentheses. *, ** and *** denote statistical significance at 5%, 1% and 0.1% respectively.

D Experimental instructions (Chapter III)

INSTRUCTIONS

Welcome and thank you for your participation in this experiment. We ask you to turn off your mobile phones and to not talk to any other participant. If you have a question, raise your hand and an assistant will come to respond to your question in person.

During this experiment, your payoff depends on your decisions and those of the other participants, as well as of luck. Below you find the rules for this experiment which determine your payoffs.

2 roles and 16 rounds

This experiment consists of **16 rounds**, each of which consists of the same sequence of decisions. This sequence is explained in detail below.
There are two kinds of roles in this experiment: **player A and player B**. At the beginning of this experiment, you are randomly assigned to one of the two roles. You will see which role is assigned to you on the first screen of the experiment. Your role remains the same throughout the experiment.

A player A always interacts with a player B. However, the pair of players changes after every round. Thus, you interact with a new player at every round.

All participants get the same information on the rules of the game, including the costs and payoffs of both roles, player A and player B.

Overview of the Sequence of Decisions in a round

Each round consists of a maximum of 4 decisions which are made consecutively. Decisions 1, 3 and 4 are made by player A, decision 2 is made by player B.

Sequence of decisions in a round (summary)

1. Player A chooses one price for action 1 and one price for action 2
2. Player B gets to know the prices chosen by player A. Then player B decides whether he wants to interact with player A. If player B does not want to interact, this round ends for both players.

If player B decides to interact with Player A :

3. Player A (but **not** player B) is informed about the type of player B. There are two possible types of player B : type 1 or type 2. Based on this, player A chooses an action for player B, either action 1 or action 2.
4. Player A charges player B one of the two prices specified at the first decision. It is not obligatory that the price charged refers to the action chosen at decision 3, player A can also charge the price for the other action.

Detailed Illustration of the Decisions and Their Consequences Regarding Payoffs

Decision 1

At decision 3, player A chooses between two actions (action 1 and action 2Each chosen action causes costs :
Action 1 costs player A **2 points** (= currency of the experiment).
Action 2 costs player A **6 points**.
Player A can charge prices for these actions from player B. At **decision 1** each Player A **sets the prices for both actions**. Only (strictly) positive integer numbers are possible, i.e., only 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 are valid prices. Note that the price for action 1 must not exceed the price for action 2.

Decision 2

Player B gets to know the prices set by player A for the two actions at decision 1. Then, player B decides whether he wants to interact with the players A.
If player B wants to interact, player A chooses an action at decision 3 and charges a price for that action at decision 4 (see below).
If player B does not want to interact, this round stops and both players receive 1.6 points for this round.

Decision 3

Before decision 3 is made (in case player B chose to interact at decision 2) a type is randomly assigned to player B. Player B can be one of the two types: **type 1 or type 2**. This type is redetermined randomly in **every new round**. With a probability of 50% player B is of type 1, and with a probability of 50% he is of type 2. Imagine

that a coin is tossed for each player B in each round. If the result is e.g. “heads”, player B is of type 1, if the result is “tails” he/she is of type 2.
Player A gets to know the type of player B before he makes his decision 3. Then player A chooses an action for player B, either action 1 or action 2.
An action is **sufficient** under the following conditions :

a) Player B is of type 1 and player A chooses action 1 or action 2.
b) Player B is of type 2 and player A chooses action 2.

An action is **insufficient** if player B is of type 2 and player A chooses action 1.

Player B receives **10 points** if the action chosen by player A is **sufficient**. Player B receives **0 points** if the action chosen by player A is **insufficient**.

Player B is **at no time** informed of his type in this round (type 1 or type 2) nor about the action (action 1 or action 2) chosen by player A.

Before player A chooses an action for player B, player B has the possibility to transfer one additional point of his payoffs to player A who is immediately informed about the outcome of this decision.
Before player A chooses an action for player B, player B has the possibility to transfer one additional point of his payoffs to player A who is immediately informed about the outcome of this decision. The transfer is only realized if the action chosen by player A is **sufficient**. If the action is insufficient, the transfer is not realized.

Decision 4
Player A charges player B a price (which he set at decision 1) for one of the actions. It is not obligatory that the price charged refers to the action chosen at decision 3, player A can also charge the price for the other action.

Payoffs

If player B chooses not to interact with player A at decision 2, both players receive **1.6 points** for this round.

If player B decides to interact at decision 2, the payoffs are as follows :

Player A receives the price (in points) which he charged at decision 4 **less** the costs for the action chosen at decision 3, and one additional point if player B has chosen to transfer an additional point to player A, and one additional point if player B has chosen to transfer an additional point to player A and player A chose a sufficient action.

For player B, the payoffs depend on **whether the action chosen by player A at decision 3 is sufficient**.

a) The action chosen by player A is **sufficient**. Player B receives 10 points less the prices charged by player A at decision 4, less 1 point if he chose to transfer an additional point to player A, less 1 point if he chose to transfer an additional point to player A.
b) The action chosen by player A is **insufficient**. Player B receives 0 points less the price charged by player A at decision 4, less 1 point if he chose to transfer an additional point to player A the transfer of the additional point is not realized.

At the beginning of the experiment you receive an initial endowment of **6 points**. With this endowment, you are able to cover losses that might occur in some rounds. Losses can also be compensated by gains in other rounds. If your total payoff after 16 rounds is negative, we ask you to perform an extra task to compensate your losses. Please notes that there is **always** a possibility to avoid losses in this experiment.

To calculate the final payoff the initial endowment and the profits of all rounds are added up.
This sum is then converted into cash using the following exchange rate::

1 point = 0.5 CHF
(i.e. 2 points = 1CHF)

Examples of decisions taken in one round

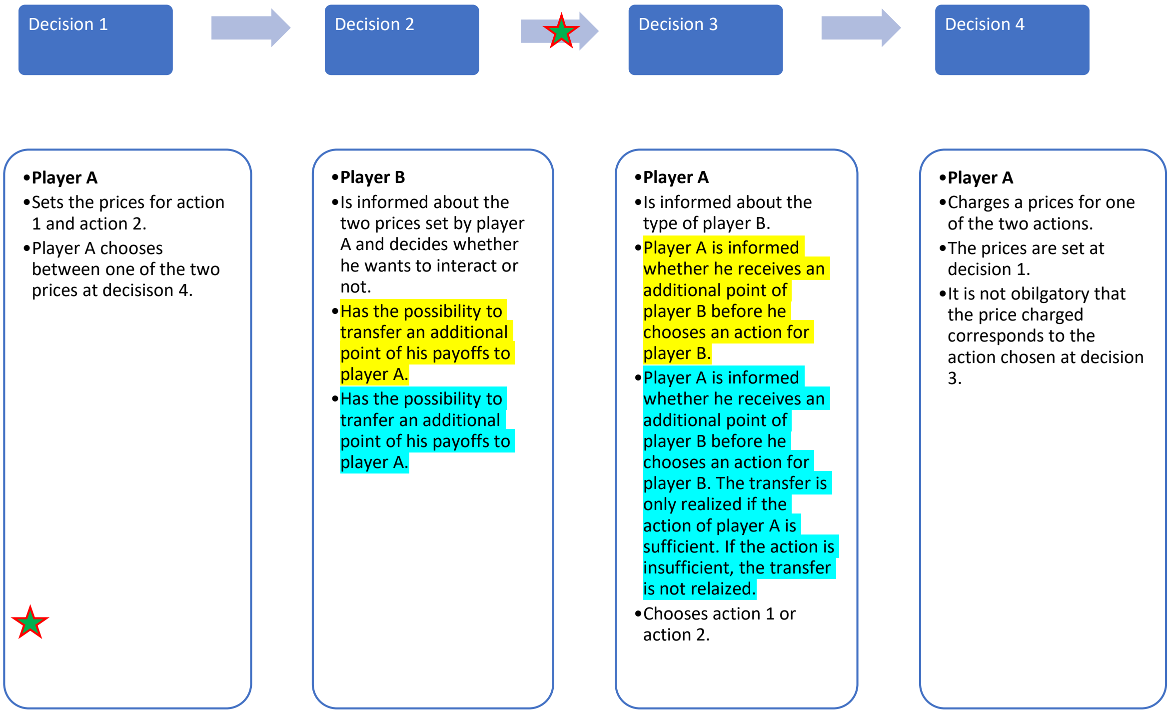
Example1 : Player A has set the following prices : price for action 1 : 2points, price for action 2 : 8 points. Player B decides not to interact at decision 2.

Payoffs for player A: 1.6
Payoffs for player B : 1.6

Example 2 :: Player A has set the following prices : price for action 1: 3 points, price for action 2: 7 points. Player B decides to interact at decision 2 and to transfer an additional point to player A and to transfer an additional point to player A. Player B is of type 2. Player A chooses action 2 and charges the price for action 2.

Payoffs for player A: 7 (the price charged by player A) – 6 (the cost of action 2)+1 (the additional point transferred by player B))+1 (the additional point transferred by player B) =1 2 2
Payoffs for player B : 10 (the value of a sufficient action) – 7 (the price paid to player A) – 1 (the additional point transferred by player B) – 1 (the additional point transferred by player B) = 3 2 2

Sequence of decisions in each round :



Decision 1

Decision 2

Decision 3

Decision 4

- **Player A**
- Sets the prices for action 1 and action 2.
- Player A chooses between one of the two prices at decisison 4.

- **Player B**
- Is informed about the two prices set by player A and decides whether he wants to interact or not.
- Has the possibility to transfer an additional point of his payoffs to player A.
- Has the possibility to tranfer an additional point of his payoffs to player A.

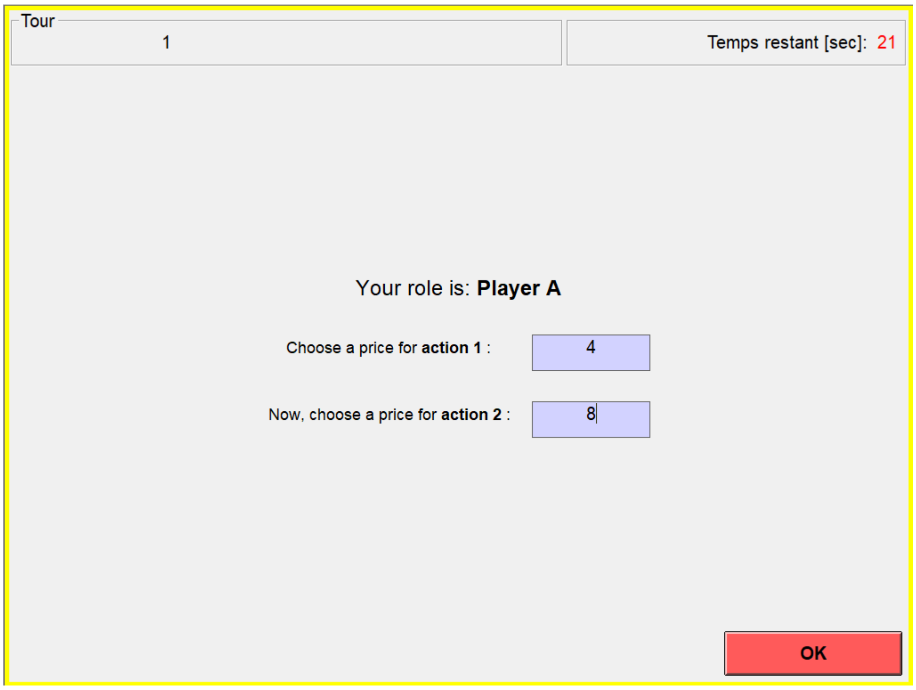
- **Player A**
- Is informed about the type of player B.
- Player A is informed whether he receives an additional point of player B before he chooses an action for player B.
- Player A is informed whether he receives an additional point of player B before he chooses an action for player B. The transfer is only realized if the action of player A is sufficient. If the action is insufficient, the transfer is not relaized.
- Chooses action 1 or action 2.

- **Player A**
- Charges a prices for one of the two actions.
- The prices are set at decision 1.
- It is not obilgatory that the price charged corresponds to the action chosen at decision 3.

If player B decides to interact, the two players continue with decisions 3 and 4. If player B decides not to interact, the round ends at decision 2.

E Screenshots (Chapter III)

Screen 1 (Player A)

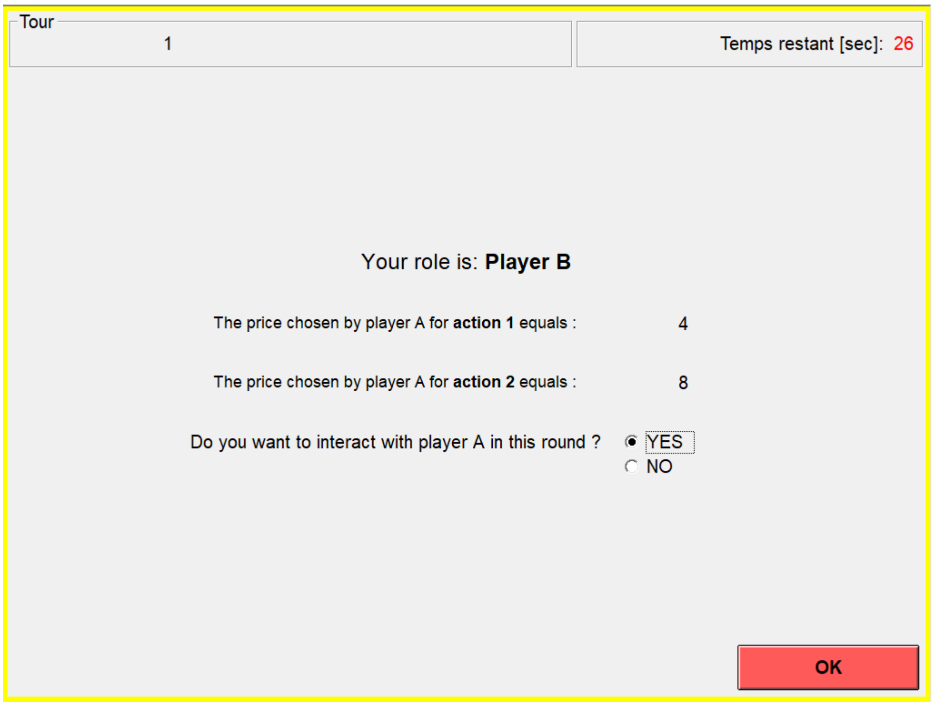


Your role is: **Player A**

Choose a price for **action 1** : 4

Now, choose a price for **action 2** : 8

Screen 2 (Player B)



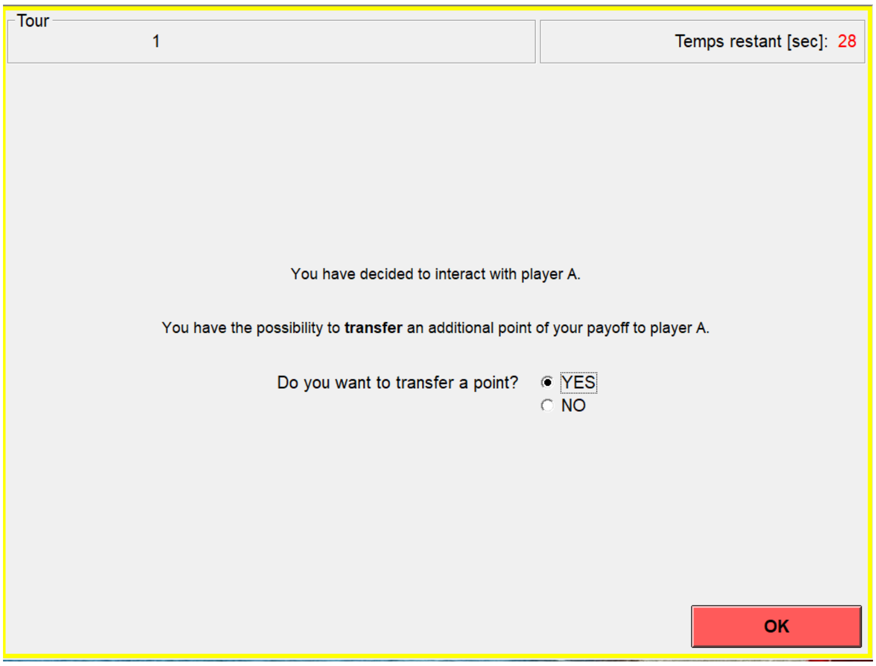
Your role is: **Player B**

The price chosen by player A for **action 1** equals : 4

The price chosen by player A for **action 2** equals : 8

Do you want to interact with player A in this round ? ● YES ○ NO

Screen 3 (Player B, only in GE and GEC)



Tour 1

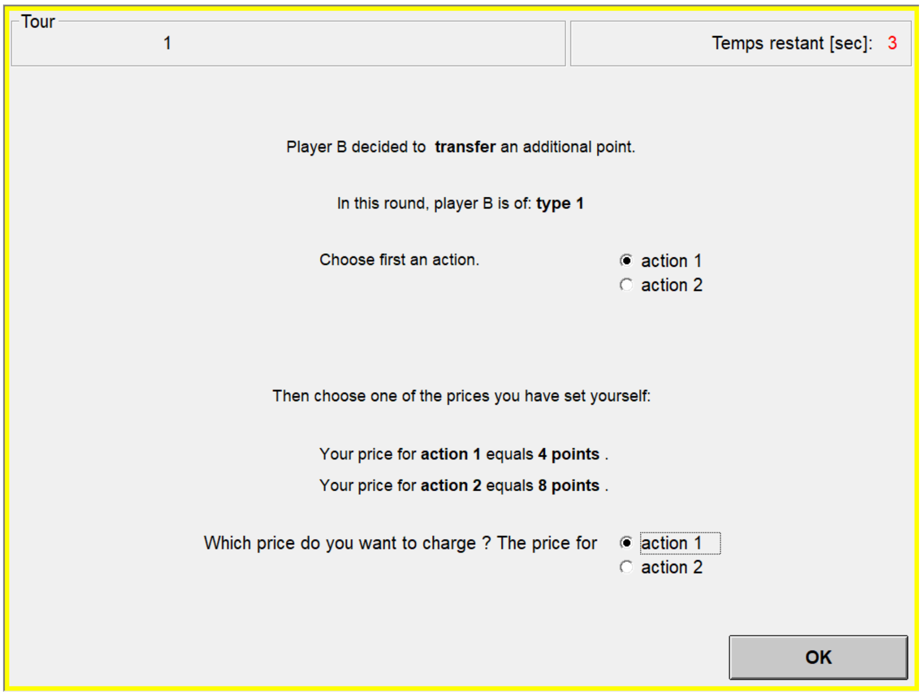
Temps restant [sec]: 28

You have decided to interact with player A.

You have the possibility to **transfer** an additional point of your payoff to player A.

Do you want to transfer a point? ● YES ○ NO

Screen 4 (Player A)



Tour 1

Temps restant [sec]: 3

Player B decided to **transfer** an additional point.

In this round, player B is of: **type 1**

Choose first an action. ● action 1 ○ action 2

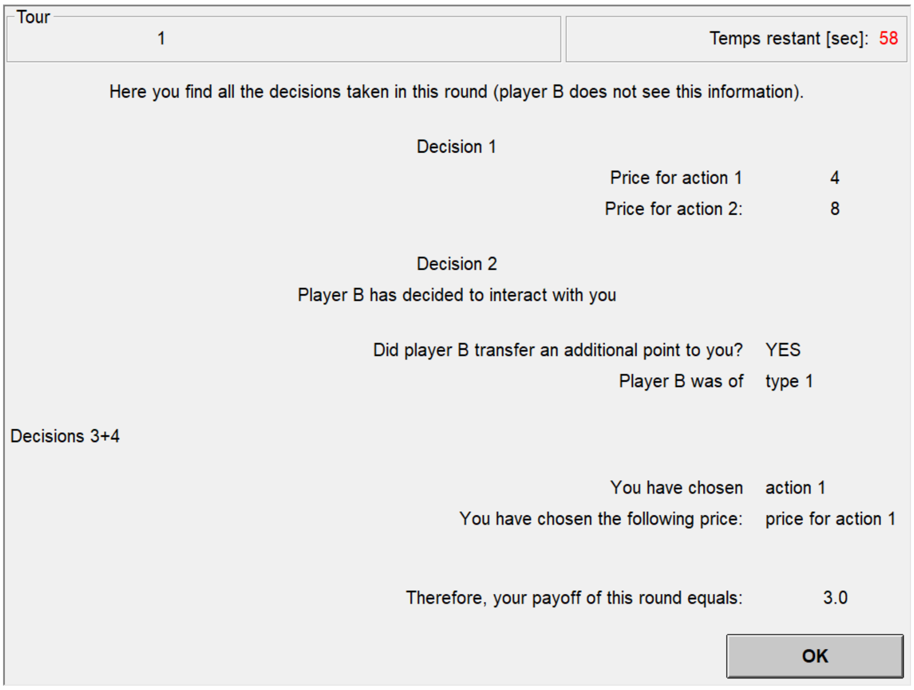
Then choose one of the prices you have set yourself:

Your price for **action 1** equals **4 points** .

Your price for **action 2** equals **8 points** .

Which price do you want to charge ? The price for ● action 1 ○ action 2

Screen 5 (Feedback forPlayer A)



Here you find all the decisions taken in this round (player B does not see this information).

Decision 1

Price for action 1 4

Price for action 2: 8

Decision 2

Player B has decided to interact with you

Did player B transfer an additional point to you? YES

Player B was of type 1

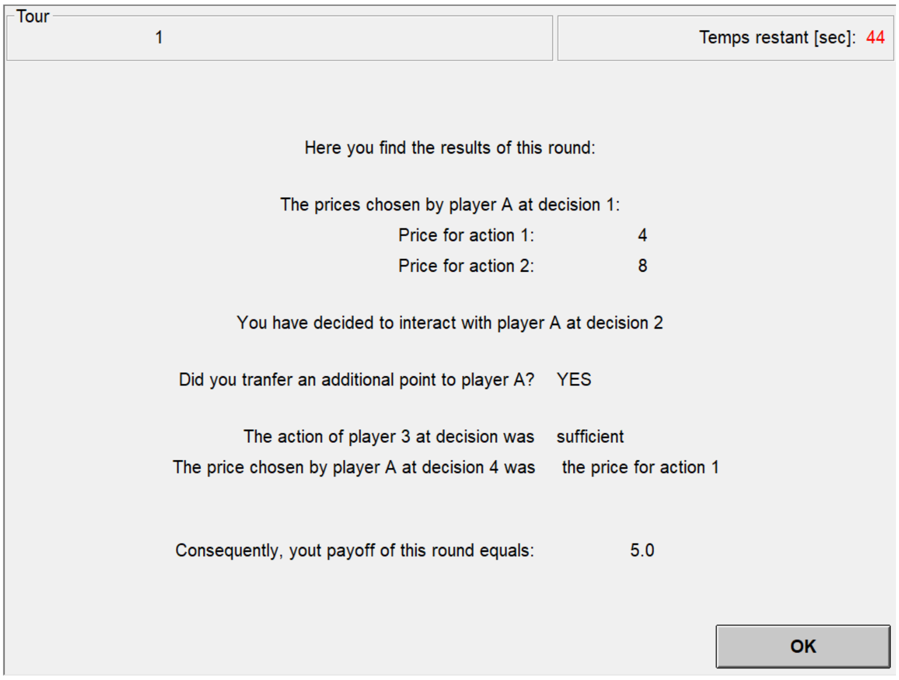
Decisions 3+4

You have chosen action 1

You have chosen the following price: price for action 1

Therefore, your payoff of this round equals: 3.0

Screen 5 (Feedback forPlayer B)



Here you find the results of this round:

The prices chosen by player A at decision 1:

Price for action 1: 4

Price for action 2: 8

You have decided to interact with player A at decision 2

Did you tranfer an additional point to player A? YES

The action of player 3 at decision was sufficient

The price chosen by player A at decision 4 was the price for action 1

Consequently, yout payoff of this round equals: 5.0

Bibliography

Abeler, J., Altmann, S., Kube, S., and Wibral, M. (2010). Gift exchange and workers' fairness concerns: When equality is unfair. *Journal of the European Economic Association*, 8(6):1299–1324.

Abrahamse, W., Steg, L., Vlek, C., and Rothengatter, T. (2005). A review of intervention studies aimed at household energy consumption. *Journal of Environmental Psychology*, 25:273–291.

Agarwal, R., Liu, C.-W., and Prasad, K. (2019). Personal research, second opinions, and the diagnostic effort of experts. *Journal of Economic Behavior & Organization*, 158:44 – 61.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, 97(4):543–569.

Alberini, A., Bareit, M., and Filippini, M. (2016). What is the effect of fuel efficiency information on car prices? Evidence from Switzerland. *The Energy Journal*, 37(3).

Aldy, J. E., Gerarden, T. D., and Sweeney, R. L. (2018). Investment versus output subsidies: Implications of alternative incentives for wind energy. NBER Working Paper No. 24378.

Allcott, H. and Greenstone, M. (2012). Is there an energy efficiency gap? *The Journal of Economic Perspectives*, 26(1):3–28.

Allcott, H. and Greenstone, M. (2017). Measuring the welfare effects of residential energy efficiency programs. NBER Working Paper No. 23386.

Allcott, H. and Knittel, C. (2019). Are consumers poorly informed about fuel economy? evidence from two experiments. *American Economic Journal: Economic Policy*, 11(1):1–37.

Allcott, H. and Sweeney, R. L. (2017). The role of sales agents in information disclosure: Evidence from a field experiment. *Management Science*, 63(1):21–39.

Allcott, H. and Wozny, N. (2014). Gasoline prices, fuel economy, and the energy paradox. *Review of Economics and Statistics*, 96(5):779–795.

Anderson, S. T. and Newell, R. G. (2004). Information programs for technology adoption: The case of energy-efficiency audits. *Resource and Energy Economics*, 26(1):27–50.

Andreoni, J. and Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636.

Andreoni, J. and Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Angelova, V. and Regner, T. (2018). Can a bonus overcome moral hazard? Experimental evidence from markets for expert services. *Journal of Economic Behavior & Organization*, 154:362 – 378.

Balafoutas, L., Beck, A., Kerschbamer, R., and Sutter, M. (2013). What drives taxi drivers? A field experiment on fraud in a market for credence goods. *Review of Economic Studies*, 80(3):876–891.

Balafoutas, L., Fornwagner, H., Kerschbamer, R., Sutter, M., and Tverdostup, M. (2020). Diagnostic uncertainty and insurance coverage in credence goods markets. *MPI Collective Goods Discussion Paper*, (2020/26).

Balafoutas, L. and Kerschbamer, R. (2020). Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions. *Journal of Behavioral and Experimental Finance*, 26:100285.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2017). Second-degree moral hazard in a real-world credence goods market. *The Economic Journal*, 127(599):1–18.

Baniamin, H. M. and Jamil, I. (2018). Institutional design for credence goods: Can the existence of financial incentive be problematic? Evidences from childbirth system of bangladesh. *International Journal of Public Administration*, 41(14):1192–1203.

Beck, A., Kerschbamer, R., Qiu, J., and Sutter, M. (2013). Shaping beliefs in experimental markets for expert services: Guilt aversion and the impact of promises and money-burning options. *Games and Economic Behavior*, 81:145–164.

Beck, A., Kerschbamer, R., Qiu, J., and Sutter, M. (2014). Car mechanics in the lab—-investigating the behavior of real experts on experimental markets for credence goods. *Journal of Economic Behavior & Organization*, 108:166–173.

Bester, H. and Dahm, M. (2017). Credence Goods, Costly Diagnosis and Subjective Evaluation. *The Economic Journal*, 128(611):1367–1394.

Blasch, J., Filippini, M., and Kumar, N. (2019). Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource and Energy Economics*, 56:39–58.

Blonz, J. A. (2019). The welfare costs of misaligned incentives: Energy inefficiency and the principal-agent problem. Unpublished manuscript.

Bollinger, B. and Gillingham, K. (2019). Learning-by-doing in solar photovoltaic installations. Unpublished manuscript.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, 90(1):166–193.

Bradley, A., Lawrence, C., and Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society B: Biological Sciences*, 285(1875):20180116.

Brounen, D., Kok, N., and Quigley, J. M. (2013). Energy literacy, awareness, and conservation behavior of residential households. *Energy Economics*, 38:42–50.

Brucal, A. and Roberts, M. J. (2019). Do energy efficiency standards hurt consumers? evidence from household appliance sales. *Journal of Environmental Economics and Management*, 96:88–107.

Burlig, F., Knittel, C. R., Rapson, D., Reguant, M., and Wolfram, C. (2020). Machine learning from schools about energy efficiency. *Journal of the Association of Environmental and Resource Economists (forthcoming)*.

Burr, C. (2016). Subsidies and investments in the solar power market. University of Colorado at Boulder Working Paper.

California Public Utilities Commision (2017). California solar initiative program handbook. San Francisco, USA.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436.

Carpenter, J. (2017). The sequencing of gift exchange: A field trial. *Journal of Economic Behavior & Organization*, 139:26 – 31.

Causholli, M., Knechel, W. R., Lin, H., and Sappington, D. E. M. (2013). Competitive procurement of auditing services with limited information. *European Accounting Review*, 22(3):573–605.

Chan, N. W. and Gillingham, K. (2015). The microeconomic theory of the rebound effect and its welfare implications. *Journal of the Association of Environmental and Resource Economists*, 2(1):133–159.

Chao, M. (2017). Demotivating incentives and motivation crowding out in charitable giving. *Proceedings of the National Academy of Sciences*, 114(28):7301–7306.

Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869.

Charness, G. and Shmidov, V. (2014). Trust and reciprocity. *Foundations and Trends in Microeconomics*, 10(3):167–207.

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

Comstock, O. and Boedecker, E. (2011). Energy and emissions in the building sector: A comparison of three policies and their combinations. *The Energy Journal*, 32:23–41.

Correia, S. (2015). Singletons, cluster-robust standard errors and fixed effects: A bad mix. *Technical Note, Duke University*.

Correia, S. (2019). Reghdfe: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. 2016 Stata Conference.

Cox, J., Friedman, D., and Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1):17–45.

Currie, J., Lin, W., and Meng, J. (2013). Social networks and externalities from gift exchange: Evidence from a field experiment. *Journal of Public Economics*, 107:19 – 30.

Currie, J., Lin, W., and Meng, J. (2014). Addressing antibiotic abuse in China: an experimental audit study. *Journal of development economics*, 110:39–51.

Darby, M. R. and Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of law and economics*, 16(1):67–88.

Davidson, C. and Steinberg, D. (2013). Evaluating the impact of third-party price reporting and other drivers on residential photovoltaic price estimates. *Energy Policy*, 62:752–761.

Davidson, C., Steinberg, D., and Margolis, R. (2015). Exploring the market for third-party-owned residential photovoltaic systems: insights from lease and power-purchase agreement contract structures and costs in california. *Environmental Research Letters*, 10(2):024006.

Davis, B., Kerschbamer, R., and Oexl, R. (2017). Is reciprocity really outcome-based? A second look at gift-exchange with random shocks. *Journal of the Economic Science Association*, 3:149–160.

Davis, L. W., Fuchs, A., and Gertler, P. (2014). Cash for coolers: Evaluating a large-scale appliance replacement program in Mexico. *American Economic Journal: Economic Policy*, 6(4):207–238.

Davis, L. W. and Metcalf, G. E. (2016). Does better information lead to better choices? Evidence from energy-efficiency labels. *Journal of the Association of Environmental and Resource Economists*, 3(3):589–625.

De Groote, O. and Verboven, F. (2019). Subsidies and time discounting in new technology adoption: Evidence from solar photovoltaic systems. *American Economic Review*, 109(6):2137–72.

Domenighetti, G., Casabianca, A., Gutzwiller, F., and Martinoli, S. (1993). Revisiting the most informed consumer of surgical services: The physician-patient. *International Journal of Technology Assessment in Health Care*, 9(4):505–513.

Dong, C., Wiser, R., and Rai, V. (2018). Incentive pass-through for residential solar systems in california. *Energy Economics*, 72:154 – 165.

Dubin, J. A., Miedema, A. K., and Chandran, R. V. (1986). Price effects of energy-efficient technologies: A study of residential demand for heating and cooling. *The RAND Journal of Economics*, 17(3):310–325.

Dulleck, U. and Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic literature*, 44(1):5–42.

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

Efficiency Valuation Organization (2018). International performance measurement and verification protocol (ipmvp). Efficiency Valuation Organization, Washington, D.C., U.S.A. Available at: https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.

Emons, W. (1997). Credence goods and fraudulent experts. *The RAND Journal of Economics*, pages 107–119.

Englmaier, F. and Leider, S. (2020). Managerial payoff and gift-exchange in the field. *Review of Industrial Organization*, 56(2):259–280.

European Commission (2013). Commission delegated regulation (EU) no 811/2013. Brussels, Belgium. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32013R0811.

Falk, A. (2007). Gift exchange in the field. *Econometrica*, 75(5):1501–1511.

Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and economic behavior*, 54(2):293–315.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *The American Economic Review*, 96(5):1611–1630.

Farrell, J. and Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives*, 10 (3):103–118.

Feger, F., Pavanini, N., and Radulescu, D. (2017). Welfare and redistribution in residential electricity markets with solar power. CEPR Discussion Paper No. DP12517.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics*, 108(2):437–459.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42(1):1 – 34.

Fehr, E. and List, J. A. (2004). The hidden costs and returns of incentives - trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5):743–771.

Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422:137–140.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178.

Fisher, C. (2010). Imperfect competition, consumer behavior, and the provision of fuel efficiency in light-duty vehicles. Resources for the Future Discussion Paper 10-60.

Fleiter, T., Schleich, J., and Ravivanpong, P. (2012). Adoption of energy-efficiency measures in SME's: An empirical analysis based on energy audit data from Germany. *Energy Policy*, 51:863–875.

Fowlie, M., Greenstone, M., and Wolfram, C. (2015). Are the non-monetary costs of energy efficiency investments large? Understanding low take-up of a free energy efficiency program. *American Economic Review*, 105(5):201–04.

Fowlie, M., Greenstone, M., and Wolfram, C. (2018). Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *The Quarterly Journal of Economics*, 133(3):1597–1644.

Frey, B. S. and Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review*, 87(4):746–755.

Gerarden, T. D., Newell, R. G., and Stavins, R. N. (2017). Assessing the energy-efficiency gap. *Journal of Economic Literature*, 55(4):1486–1525.

Gilbert, B., LaRiviere, J., and Novan, K. (2019). Additionality, mistakes, and energy efficiency investment. Working Papers 2019-01, Colorado School of Mines, Division of Economics and Business.

Gillingham, K., Deng, H., Wiser, R., Darghouth, N., Nemet, G., Barbose, G., Rai, V., and Dong, C. (2016a). Deconstructing solar photovoltaic pricing. *The Energy Journal*, 37(3).

Gillingham, K. and Palmer, K. (2014). Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy*, 81:18–38.

Gillingham, K., Rapson, D., and Wagner, G. (2016b). The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy*, 10(1):68–88.

Gillingham, K. and Tsvetanov, T. (2018). Nudging energy efficiency audits: Evidence from a field experiment. *Journal of Environmental Economics and Management*, 90:303–316.

Giraudet, L.-G. (2020). Energy efficiency as a credence good: A review of informational barriers to energy savings in the building sector. *Energy Economics*, 87:104698.

Giraudet, L.-G., Houde, S., and Maher, J. (2018). Moral hazard and the energy efficiency gap: Theory and evidence. *Journal of the Association of Environmental and Resource Economists*, 5(4):755–790.

Globus-Harris, I. (2020). Waiting periods as a screening mechanism for environmental subsidies. *Journal of the Association of Environmental and Resource Economists (forthcoming)*.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1):384–394.

Gneezy, U. and Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics*, 115(3):791–810.

Godager, G. and Wiesen, D. (2013). Profit or patients' health benefit? Exploring the heterogeneity in physician altruism. *Journal of Health Economics*, 32:1105–1116.

Goeschl, T. (2019). Cold case: The forensic economics of energy efficiency labels for domestic refrigeration appliances. *Energy Economics*, 84:104468.

Gottschalk, F., Mimra, W., and Waibel, C. (2020). Health Services as Credence Goods: A Field Experiment. *The Economic Journal (forthcoming)*.

Greiner, B., Zhang, L., and Tang, C. (2017). Separation of prescription and treatment in health care markets: A laboratory experiment. *Health Economics*, 26(S3):21–35.

Gruber, J., Kim, J., and Mayzlin, D. (1999). Physician fees and procedure intensity: The case of Cisarean delivery. *Journal of Health Economics*, 18:473–490.

Gruber, J. and Owings, M. (1996). Physician financial incentives and Cesarean section delivery. *RAND Journal of Economics*, 27:99–123.

Güth, W., Huck, S., and Ockenfels, P. (1996). Two-level ultimatum bargaining with incomplete information: An experimental study. *The Economic Journal*, 106(436):593–604.

Hannan, R. L., Kagel, J. H., and Moser, D. V. (2002). Partial gift exchange in an experimental labor market: Impact of subject population differences, productivity differences, and effort requests on behavior. *Journal of Labor Economics*, 20(4):923–951.

Hennig-Schmidt, H., Sadrieh, A., and Rockenbach, B. (2010). In search of workers' real effort reciprocity - A field and a laboratory experiment. *Journal of the European Economic Association*, 8(4):817–837.

Hennig-Schmidt, H., Selten, R., and Wiesen, D. (2011). How payment systems affect physicians' provision behaviour - an experimental investigation. *Journal of Health Economics*, 30(4):637 – 646.

Heutel, G. and Muehlegger, E. (2015). Consumer learning and hybrid vehicle adoption. *Environmental and Resource Economics*, 61 (1):125–161.

Hoppe, E. and Schmitz, P. (2018). Hidden action and outcome contractibility: An experimental test of moral hazard theory. *Games and Economic Behavior*, 109:544–564.

Houde, S. (2018). Bunching with the stars: How firms respond to environmental certification. Center of Economic Research at ETH Zurich Working Paper 18/292.

Houde, S. and Spurlock, C. A. (2015). Do energy efficiency standards improve quality? evidence from a revealed preference approach. Lawrence Berkeley National Laboratory Report, LBNL-182701.

Huck, S., Lünser, G., Spitzer, F., and Tyran, J.-R. (2016). Medical insurance and free choice of physician shape patient overtreatment: A laboratory experiment. *Journal of Economic Behavior & Organization*, 131:78–105.

Hughes, J. E. and Podolefsky, M. (2015). Getting green with solar subsidies: Evidence from the california solar initiative. *Journal of the Association of Environmental and Resource Economists*, 2(2):235–275.

Iizuka, T. (2007). Experts' agency problem: Evidence from the prescription drug market in Japan. *RAND Journal of Economics*, 38:844–862.

Iizuka, T. (2012). Physician agency and adoption of generic pharmaceuticals. *American Economic Review*, 102:2826–2858.

International Energy Agency (2016). World energy outlook 2016. Paris, France.

Jacobsen, G. D. (2015). Do energy prices influence investment in energy efficiency? Evidence from energy star appliances. *Journal of Environmental Economics and Management*, 74:94–106.

Jaffe, A. B. and Stavins, R. N. (1994a). The energy-efficiency gap what does it mean? *Energy Policy*, 22(10):804–810.

Jaffe, A. B. and Stavins, R. N. (1994b). Energy-efficiency investments and public policy. *The Energy Journal*, 15(2):43–65.

Jessoe, K. and Rapson, D. (2014). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review*, 104 (4):1417–1438.

Joskow, P. L. and Marron, D. B. (1992). What does a negawatt really cost? Evidence from utility conservation programs. *The Energy Journal*, 13(4):41–74.

Kandul, S., Lanz, B., and Reins, E. (2020). Reciprocity and gift exchange in markets for credence goods. University of Neuchâtel, IRENE Working Paper 20-09, Institute of Economic Research.

Kerschbamer, R., Neururer, D., and Sutter, M. (2016). Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proceedings of the National Academy of Sciences*, 113(27):7454–7458.

Kerschbamer, R. and Sutter, M. (2017). The economics of credence goods – a survey of recent lab and field experiments. *CESifo Economic Studies*, 63(1):1–23.

Kerschbamer, R., Sutter, M., and Dulleck, U. (2017). How social preferences shape incentives in (experimental) markets for credence goods. *The Economic Journal*, 127(600):393–416.

Kirchler, M. and Palan, S. (2018). Immaterial and monetary gifts in economic transactions: Evidence from the field. *Experimental Economics*, 21:205–230.

Klinke, S. (2018). The determinants for adoption of energy supply contracting: Empirical evidence from the swiss market. *Energy Policy*, 118:221–231.

Knechel, W. R. (2013). Do auditing standards matter? *Current Issues in Auditing*, 7(2):A1–A16.

Kube, S., Maréchal, M. A., and Puppe, C. (2012). The currency of reciprocity: Gift exchange in the workplace. *American Economic Review*, 102(4):1644–62.

Lanz, B. and Reins, E. (2021). Asymmetric information on the market for energy efficiency: Insights from the credence goods literature. *The Energy Journal*, 42(4).

Lilly, P. and Simons, G. (2006). California's self-generation incentive program nonresidential pv systems: Measured system performance and actual costs. In *ASME Power Conference*, volume 42053, pages 667–673.

Lu, F. (2014). Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in china. *Journal of Development Economics*, 106:156–167.

Malmendier, U. and Schmidt, K. M. (2017). You owe me. *American Economic Review*, 107(2):493–526.

Mellström, C. and Johannesson, M. (2008). Crowding out in blood donation: Was Titmuss right? *Journal of the European Economic Association*, 6(4):845–863.

Milgrom, P. (2008). What the seller won't tell you: Persuasion and disclosure in markets. *Journal of Economic Perspectives*, 22(2):115–131.

Mimra, W., Rasch, A., and Waibel, C. (2016a). Price competition and reputation in credence goods markets: Experimental evidence. *Games and Economic Behavior*, 100:337–352.

Mimra, W., Rasch, A., and Waibel, C. (2016b). Second opinions in markets for expert services: Experimental evidence. *Journal of Economic Behavior & Organization*, 131:106–125.

Montinari, N., Nicolò, A., and Oexl, R. (2016). The gift of being chosen. *Experimental Economics*, 19(2):460–479.

Nadel, S. and Keating, K. (1991). Engineering estimates vs. impact evaluation results: How do they compare and why? Research Report U915, American Council for an Energy-Efficient Economy. Available at: http://www.aceee.org/research-report/u915.

Newell, R. G. and Siikamäki, J. (2014). Nudging energy efficiency behavior: The role of information labels. *Journal of the Association of Environmental and Resource Economists*, 1(4):555–598.

Newman, G. E. and Jeremy Shen, Y. (2012). The counterintuitive effects of thank-you gifts on charitable giving. *Journal of Economic Psychology*, 33(5):973 – 983.

O'Neill, M. (2009). Public confidence in charitable nonprofits. *Nonprofit and Voluntary Sector Quarterly*, 38(2):237–269.

Palm, A. (2017). Peer effects in residential solar photovoltaics adoption—a mixed methods study of swedish users. *Energy Research & Social Science*, 26:1–10.

Plambeck, E. L. and Taylor, T. A. (2019). Testing by competitors in enforcement of product standards. *Management Science*, 65(4):1455–1947.

Pless, J. and van Benthem, A. A. (2019). Pass-through as a test for market power: An application to solar subsidies. *American Economic Journal: Applied Economics*, 11(4):367–401.

Podolefsky, M. (2013). Tax evasion and subsidy pass-through under the solar investment tax credit. University of Colorado at Boulder Working Paper 13-05.

Rasch, A. and Waibel, C. (2017). What drives fraud in a credence goods market? Evidence from a field study. *Oxford Bulletin of Economics and Statistics*, 80(3):605–624.

Rottenberg, S. (1965). Liability in law and economics. *The American Economic Review*, 55(1/2):107–114.

Rubin, J. and Sheremeta, R. (2015). Principal-agent settings with random shocks. *Management Science*, 62 (4):905–1224.

Sallee, J. (2014). Rational inattention and energy efficiency. *Journal of Law and Economics*, 57(3):781–820.

Salzman, A. (2013). Dark clouds over solarcity. *Barron's Magazine*. Available at https://www.barrons.com/articles/SB50001424052748704719204579025283044181654

Schneider, H. S. (2012). Agency problems and reputation in expert services: Evidence from auto repair. *The Journal of Industrial Economics*, 60(3):406–433.

Soraperra, I., Suvorov, A., de Ven, J. V., and Villeval, M. C. (2019). Doing bad to look good: Negative consequences of image concerns on prosocial behavior. *Revue Economique*, 70(6):945–966.

Sorrell, S. (2004). Understanding barriers to energy efficiency. In Sorrell, S., O'Malley, E., Schleich, J., and Scott, S., editors, *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment*, chapter 2, pages 25–93. Edward Elger, Cheltenham.

Sorrell, S. (2007). The economics of energy service contracts. *Energy Policy*, 35 (1):507–521.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8(1):321–340.

Tang, J. J. (2020). Individual heterogeneity and cultural attitudes in credence goods provision. *European Economic Review*, 126:103442.

Tietenberg, T. (2009). Reflections-energy efficiency policy: Pipe dream or pipeline to the future? *Review of Environmental Economics and Policy*, 3(2):304–320.

Trabish, H. K. (2013). Why treasury is investigating solarcity and solar third-party funds. Greentech Media. Available at https://www.greentechmedia.com/articles/read/why-treasury-is-investigating-solarcity-and-solar-third-party-funds.

U.S. Department of Energy (2012). National energy audit tool (neat). Weatherization Assistance Program, Washington D.C., U.S.A. Available at: https://www.buildingenergysoftwaretools.com/software/national-energy-audit-tool-neat.

U.S. Department of Energy (2018). Energy saver program. Office of Energy Efficiency and Renewable Energy, Washington D.C., U.S.A. Available at: https://www.energy.gov/energysaver/energy-saver.

U.S. Environmental Protection Agency (2017). Notice of violation (18 september 2015): sent by epa to volkswagen-group of america, inc. U.S. Environmental Protection Agency, Washington D.C., U.S.A. Available at: https://www.epa.gov/sites/default/files/2015-10/documents/vw-nov-caa-09-18-15.pdf.

U.S. Environmental Protection Agency (2018). Emission standards reference guide. U.S. Environmental Protection Agency, Washington D.C., U.S.A. Available at: https://www.epa.gov/emission-standards-reference-guide.

U.S. Federal Trade Commission (1979). Federal register: 44 fed. reg. 66466. monday, 1979. National Archives and Records Administration, Washington D.C., U.S.A. Available at: https://www.federalregister.gov/documents/2020/04/10/2020-06960/energy-labeling-rule.

Wolske, K. S., Gillingham, K. T., and Schultz, P. W. (2020). Peer influence on household energy behaviours. *Nature Energy*, 5(3):202–212.