Double Moral Hazard and the Energy Efficiency Gap

Louis-Gaëtan Giraudet*, Sébastien Houde†

First draft: June 27th, 2012
This draft: May 2014

Abstract

We investigate how moral hazard problems can cause sub-optimal investment in energy efficiency, a phenomenon known as the *energy efficiency gap*. We argue that such problems are likely to be important for home energy retrofits, where both the seller and the buyer can take hidden actions. The retrofit contractor may cut on the quality of installation to save costs, while the homeowner may 'rebound', that is, increase her use of energy service when provided with higher energy efficiency. We first formalize the double moral hazard described above and examine how the resulting energy efficiency gap can be reduced through policy intervention. We find that minimum quality standards outperform energy-savings insurance, which is incomplete in equilibrium. We then calibrate the model to the U.S. home insulation context. Numerically, moral hazard problems are consistent with homeowners investing with implicit discount rates in the 20-30% range. The welfare gains from undoing moral hazard are several times as large as the costs of quality audits. They are also about one order of magnitude larger than those from internalizing carbon dioxide externalities associated with the use of natural gas for space heating.

JEL: D86,Q41
Keywords: Energy efficiency gap, Moral hazard, Energy-savings insurance, Minimum quality standard, Rebound effect

*Centre International de Recherche sur l'Environnement et le Développement (CIRED), Ecole des Ponts ParisTech, and Precourt Energy Efficiency Center (PEEC), Stanford University. Email: giraudet@centre-cired.fr.

†Department of Agricultural and Resource Economics, University of Maryland. Email: shoude@umd.edu.

1. Introduction

The rationale for government interventions promoting energy efficiency has been debated for more than three decades. At the heart of the debate are empirical studies finding that people apply abnormally high discount rates to energy efficiency investment decisions (see Train (1985) for an early review). This suggests that some investment opportunities that are privately profitable - even if energy-use externalities are not internalized - are not undertaken. This empirical fact is known as the energy efficiency gap. Jaffe and Stavins (1994) were the first to conceptualize this debate by emphasizing the difference between market failure and "non-market failure" explanations of the energy efficiency gap, and to argue that only market failures in energy efficiency markets could justify government interventions.

In this paper, we aim to contribute to this debate by drawing attention to one market failure which, to our knowledge, has been overlooked in the literature: moral hazard in the provision of quality in energy efficiency investment[1]. We also shed light on little discussed policy remedies such as energy-savings insurance and the certification of energy efficiency professionals.

Many energy efficiency technologies are considered to be credence goods, the performance of which never gets perfectly known to the buyer (Sorrell, 2004). This characteristic is conducive to a variety of information problems, which have long been suspected to be the main source of market failures in energy efficiency markets (Howarth and Andersson, 1993; Huntington *et al.*, 1994). This is especially true in the building sector. Technological complexity may cause a general lack of understanding about energy saving opportunities. However, evaluations of energy audits find that consumers respond less to information provision than to price signals, suggesting that the knowledge gap may be small (Palmer *et al.*, 2013; Frondel and Vance, 2013). Still, information may be comprehensible but asymmetrically distributed. One set of works have examined information asymmetries in home sales, in which the seller is supposedly more informed than the buyer about the energy efficiency performance of the dwelling. The empirical results are mixed. Some authors find that energy efficiency tends to be capitalized into sale prices (Brounen and Kok, 2011; Kok and Kahn, 2012) while others find opposite results, more consistent with the hidden information model (Myers, 2014). Another set of works have examined information asymmetries in rental housing, in which the landlord is supposedly more informed than the tenant about the energy efficiency performance of the dwelling. They find that higher energy efficiency does not lead to higher

[1]For instance, moral hazard problems are not discussed in the most recent and exhaustive literature reviews on the energy efficiency gap (Gillingham *et al.*, 2009, Allcott and Greenstone, 2012).

rents and that rented dwellings are significantly less energy efficient than owner-occupied ones (Davis, 2010; Gillingham *et al.*, 2012).

The information asymmetries we consider here are related, but different in nature. We examine energy efficiency projects in which a contractor may cut on the quality of installation[2] to save costs, while the buyer may 'rebound', *i.e.*, increase her use of energy service when provided with higher energy efficiency. We refer to this problem as double moral hazard. Our focus is on the supply side of energy efficiency markets; we thus extend an analysis of information asymmetries so far confined to building ownership and occupancy.

Our contribution is threefold. We first formalize how moral hazard in the provision of quality leads to an energy efficiency gap. We then investigate policy tools that could be used to address this market failure. In the building sector, firms could offer energy-saving insurance (Mills, 2003). However, we show that due to unobservable homeowner's behavioral response, a complete insurance contract is not optimal. We also examine professional certification in the form of minimum quality standards. These policy tools can aptly complement more studied energy efficiency policy tools such as energy taxes (Allcott *et al.*, 2014), energy efficiency subsidies (Boomhower and Davis, 2014; Ito, 2013), energy efficiency labels (Houde, 2014), building codes (Aroonruengsawat *et al.*, 2012; Jacobsen and Kotchen, 2013) and information provision (Jessoe and Rapson, 2014). We finally quantify the size of the energy efficiency gap due to moral hazard in the U.S. insulation market, using data from the U.S. 2009 Residential Energy Consumption Survey (RECS). We find that moral hazard problems leads to implicit discount rates in the 20-30% range, instead of a 7% rate assumed in the absence of this market failure. As a result, energy efficiency levels are too low, participation to investment is deterred and welfare losses are substantial. To put numbers in perspective, we find that undoing moral hazard problems bring welfare gains several times as large as the costs of quality audits. They are also approximately one order of magnitude larger than those from internalizing carbon dioxide externalities associated with natural gas use. Assuming some heterogeneity with respect to energy service valuation in the population, we find uniform standards and insurance that yield only negligible welfare losses due to contractors or the Government being unable to screen consumer types.

The outline of the paper is as follows. Section 2 introduces the model. Section 3 examines quality standards and energy-savings insurance as policy solutions to the moral hazard

[2]Such a quality shortfall may materialize as either inefficient labor or capital input. For instance, an insulation contractor may omit to fill wall cavities before installing insulation panels and/or install insulation panels of a low grade.

problem. Section 4 provides numerical estimates of the size of the energy efficiency gap due to moral hazard problems in the U.S. home insulation market and assesses different policy instruments. Section 5 discusses the modeling assumptions and predictions. Section 6 concludes.

2. **Energy efficiency investment with double moral hazard**

Our model builds upon the double moral hazard model of Cooper and Ross (1985). Investments in energy retrofits, which typically involve hidden actions from both the homeowner and the contractor, are considered as a canonical example. Other occupancy status that merely give rise to one-sided moral hazard can be viewed as special cases of this general model; they are occasionally discussed in the text.

2.1. **Set-up**

A homeowner consumes energy for space heating. This energy service s, measured in indoor temperature, provides her with value $V(s)$, multiplied by a taste parameter $\theta > 0$ representing heterogeneity across consumers in the valuation of energy service. The homeowner expects to pay energy bill $pE^0(s)$, where $E^0(\cdot)$ is energy use and p is the price of energy, assumed to be constant over the period considered. Energy use is a random variable influenced by idiosyncratic factors, such as weather conditions and the architectural characteristics of the house. For simplicity, we use a deterministic framework; utility is quasi-linear and there is no risk aversion. The homeowner consumes the energy service at the optimal level s^0_θ that maximizes expected utility $U^0(\theta, s)$ considered over a technical lifetime of l years with some discount factor δ:

$$U^0(\theta, s) \equiv \sum_{t=1}^{l} \left[\theta V(s) - pE^0(s)\right] \delta^t \tag{1}$$

The homeowner invests in retrofits to reduce her energy bill. In this setting, the homeowner is the principal and the contractor is the agent. Energy use after investment E is reported on homeowner's energy statement. Hence, it is common knowledge to both parties. Yet each one can take hidden actions s and q to influence it.

The homeowner chooses a level of energy service s. This action is unobserved to the contractor and a higher energy service will induce a higher expected energy use. Likewise, the contractor provides a certain quality q in installing energy efficient equipments. Variable q can be measured as the number of hours worked by the installers employed by the contractor and is bounded below by a minimum input q_{min}. Unlike on other amenities, the impact of this action on the energy efficiency performance cannot be fully assessed by the homeowner. The only thing that is known to both parties is that a higher quality of installation lowers expected energy use.

The homeowner considers future discounted benefits with expected utility $U(\theta, s, q)$, where $T > 0$ is the upfront cost of the retrofit and ϵ denotes some fixed non-energy benefits associated with it, net of the inconvenience costs generated by the investment:

$$U(\theta, s, q) \equiv \sum_{t=1}^{l} [\theta V(s) - pE(s, q)]\, \delta^t - T + \epsilon \tag{2}$$

Firms are homogenous in the industry. The profit of a representative contractor is the revenue from the sale minus the cost of the quality provided:

$$\Pi(q) \equiv T - C(q) \tag{3}$$

The following assumptions hold (subscripts denote partial derivatives):

Assumption 1: Technology.
(i) At constant consumer behavior s, investment reduces energy use: $\forall q \geq q_{min}\, E(s, q) < E^0(s)$
(ii) Contracting parties' actions have opposite effects: $E^0_s > 0$, $E_s > 0$ and $E_q < 0$
(iii) Energy savings exhibit decreasing returns: $-E^0_{ss} \leq 0$, $-E_{ss} \leq 0$ and $-E_{qq} \leq 0$
(iv) Contracting parties' actions are substitutes: $E_{qs} < 0$ and $E_s < E^0_s$
(v) Non-energy benefits are not sufficient to motivate investment: $\epsilon \leq C(q_{min})$

Assumption 2: Behavior. *Contracting parties are (i) value-maximizers, (ii) risk-neutral and (iii) have twice differentiable, concave value functions: $V'(\cdot) > 0$,$V''(\cdot) \leq 0$ and $-C'(\cdot) < 0$, $-C''(\cdot) \leq 0$*

Assumption 3: Market. *The industry is competitive with free entry:* $\Pi(q) = 0$.
Corollary: $T \equiv C(q)$.

Assumption 1 is mild: The energy service has a convex effect on expected energy use, and quality has diminishing returns on expected energy savings. Moreover, both factors impede each other: The marginal increase in expected energy savings due to increased quality is larger when underlying energy service is high (*e.g.*, a house heated in a cold climate) rather than low (*e.g.*, a house heated in a warm climate). Reciprocally, the marginal increase in expected energy use due to increased energy service is lower when the quality installed is high rather than low.

Assumptions 2 and 3 are meant to be as standard as possible, in order to isolate the moral hazard problem from possibly interacting market failures. Their generality is discussed in Section 5.

2.2. Social ***versus*** private optimum

We will consider two equilibrium outcomes: a social (hereafter cooperative) optimum c and a private (hereafter non-cooperative) optimum nc. For any equilibrium situation $j \in \{c, nc\}$, the agreement between the homeowner and the contractor is a two-stage game that is solved backward. In the first stage, the homeowner of type θ invests if the net present value $NPV^j(\theta)$ of the investment is positive, given her beliefs about her future optimal energy service s_θ^j and the optimal quality offered to her by the contractor q_θ^j:

$$NPV^j(\theta) \equiv U(\theta, s_\theta^j, q_\theta^j) - U_0(\theta, s_\theta^0) \geq 0 \tag{4}$$

In the second stage, both agents determine their own action given their belief about the other party's action. We focus hereafter on this second stage, for a participating consumer of type θ.

Under perfect information, the contract between the two parties is set cooperatively so as to maximize joint expected surplus, subject to boundary conditions $s \geq s_{min}$ and $q \geq q_{min}$. The optimal actions s_θ^c and q_θ^c that solve the first-order conditions for maximization[3] below will be such that their marginal benefit (in terms of value to the consumer and cost savings

[3]Throughout the paper, the objective functions are well-behaved and the first-order conditions discussed are necessary and sufficient for maximization.

to the firm) equates their marginal effect on consumer's expected energy bill:

$$\forall t \quad \theta V' \leq pE_s \quad \text{with equality if} \quad s^c_\theta > s_{min} \tag{5}$$

$$C' \geq -\sum_{t=1}^{l} pE_q\delta^t \quad \text{with equality if} \quad q^c_\theta > q_{min} \tag{6}$$

The cooperative optimum (s^c_θ, q^c_θ) can be characterized as a reaction function equilibrium. Assuming interior solutions and applying the Implicit Function Theorem to the first-order conditions, we find that the reaction functions $s^*_\theta(q)$ and $q^*(s)$ are strictly increasing:

$$\forall t \quad \frac{\mathrm{d}s^*_\theta}{\mathrm{d}q} = \frac{pE_{qs}}{\theta V'' - pE_{ss}} > 0 \tag{7}$$

$$\frac{\mathrm{d}q^*}{\mathrm{d}s} = \frac{-\sum_{t=1}^{l} pE_{sq}\delta^t}{C'' + \sum_{t=1}^{l} pE_{qq}\delta^t} > 0 \tag{8}$$

Now if information is imperfect, the agreement is no longer cooperative. Both parties maximize their private expected value, given their beliefs about other party's action and subject to boundary conditions $s \geq s_{min}$ and $q \geq q_{min}$. While this yields the same reaction function as in the cooperative agreement $s^*_\theta(q)$ for the consumer, this does not hold for the contractor. He does not internalize the expected benefits that his action delivers to the homeowner and simply chooses the level of quality q^{nc} that minimizes his cost:

$$\forall s \quad q^{nc}(s) = \underset{q \geq q_{min}}{\arg\min}\, C(q) = q_{min} \tag{9}$$

Proposition 1. *For a participating consumer of given type θ:*
(i) the private, non-cooperative equilibrium $(s^{nc}_\theta, q^{nc}_\theta)$ exists and is unique
(ii) the social, cooperative equilibrium (s^c_θ, q^c_θ) exists and is unique if and only if:

$$\frac{\mathrm{d}q}{\mathrm{d}s^*_\theta} > \frac{\mathrm{d}q^*}{\mathrm{d}s} \tag{10}$$

proof: (i) The private equilibrium is uniquely defined as $(s^*_\theta(q_{min}), q_{min})$. (ii) Likewise, if at least one agent has his or her optimal cooperative action at corner, then the social equilibrium is uniquely defined. If optimal actions are interior for both agents, condition (10) implies that the composite function $s^*_\theta(q^*(s))$ defined for all $s \geq s_{min}$ is a contraction mapping. Hence, by the Banach fixed-point theorem, it admits a unique fixed point.

The following proposition states that the two equilibria will involve unambiguous locations:

Proposition 2. *Assuming condition (10) holds, a participating consumer of given type θ:*
(i) is offered a higher level of quality in the social optimum: $q^c_\theta \geq q^{nc}_\theta$
(ii) sets her energy service at a higher level in the social optimum: $s^c_\theta \geq s^{nc}_\theta > s^0_\theta$
(iii) faces a higher net present value in the social optimum: $NPV^c(\theta) \geq NPV^{nc}(\theta)$

proof: (i) For a given θ, $q^c_\theta \geq q_{min} = q^{nc}_\theta$. (ii) Since $s^*_\theta(\cdot)$ is increasing, $s^c_\theta = s^*_\theta(q^c_\theta) \geq s^*_\theta(q^{nc}_\theta) = s^{nc}_\theta$. For all s, $E^0_s > E_s$ implies $U_s > U^0_s$. Therefore, assuming interior solutions: $U^0_s|_{s^0_\theta} = 0 = U_s|_{s^{nc}_\theta} > U^0_s|_{s^{nc}_\theta}$. Since U^0 is concave in s, U^0_s is decreasing in s and $s^{nc}_\theta > s^0_\theta$. (iii) Comparing net present values $NPV^c(\cdot)$ and $NPV^{nc}(\cdot)$ is equivalent to comparing the expected utility functions after investment $U(\theta, s^c_\theta, q^c_\theta)$ and $U(\theta, s^{nc}_\theta, q^{nc}_\theta)$. Under the assumption of perfect competition, the expected utility after investment is equivalent to the joint expected surplus. Therefore, the net present value of investment is maximized in the social outcome: $NPV^c(\theta) \geq NPV^{nc}(\theta)$.

Recall from Assumption (1ii) that q and s have an opposite effect on $E(s, q)$. Hence, if both inputs increase simultaneously, as is the case when the parties move from the private to the social optimum, the decrease in energy use due to the increase in energy efficiency quality is partly offset by the increase in energy service. This phenomenon is known as the 'rebound effect'. To the extreme, it can 'backfire', *i.e.*, be such that energy use increases after energy efficiency investments. This case cannot be ruled out from our analysis, as $E(s^c_\theta, q^c_\theta)$, $E(s^{nc}_\theta, q^{nc}_\theta)$ and $E(s^0_\theta)$ cannot be compared unambiguously.

We shall now make a distinction between two types of backfire rebound effect, which will prove useful later in the analysis.

Definition 1: Genuine backfire rebound effect. *A genuine backfire rebound effect occurs if energy use after investment is larger than before investment: $s > s^0$ and $E(s, q) > E^0(s^0)$*

Definition 2: Relative backfire rebound effect. *A relative backfire rebound effect occurs between two investment options H and L if energy use after investment is larger in the more energy efficient option H: $q^H > q^L$, $s^H > s^L$ and $E(s^H, q^H) > E(s^L, q^L)$*

2.3. Consumer heterogeneity and aggregate welfare

We now turn to a continuum of consumers of mass 1. Consumers are assumed to all live in a similar dwelling and only differ with respect to their preference for energy service θ. The higher the type θ, the higher the demand for energy service, hence the higher the quality offered by a cooperative firm; in contrast, the quality offered by a non-cooperative firm remains at minimum. This proposition is demonstrated in Appendix A.

For any equilibrium situation $j \in \{c, nc\}$, we have, by the Envelope Theorem:

$$\frac{\mathrm{d}NPV^j}{\mathrm{d}\theta} = \sum_{t=1}^{l} \left[V(s_\theta^j) - V(s_\theta^0)\right] \delta^t \tag{11}$$

As $V(\cdot)$ is increasing and $\forall\theta\ s_\theta^j > s_\theta^0$, equation 11 means that the net present value of investment strictly increases with θ. Hence, if a cutoff type θ_0^j such that $NPV^j(\theta_0^j) = 0$ exists, it is unique. In what follows, we are interested in this most relevant case; alternative cases are discussed in Appendix B. Assuming that $F(\cdot)$ is the cumulative distribution function of θ, participation to investment N^j is given by:

$$N^j \equiv 1 - F(\theta_0^j) \tag{12}$$

Finally, aggregate social welfare is the sum of utility before investment for those consumers who do not invest (*i.e.*, $\theta \in [0, \theta_0)$), plus the utility after investment for those who do invest (*i.e.*, $\theta \geq \theta_0$):

$$W^j \equiv \int_0^{\theta_0^j} U^0(\theta, s_\theta^0)\mathrm{d}F(\theta) + \int_{\theta_0^j}^{+\infty} U(\theta, s_\theta^j, q_\theta^j)\mathrm{d}F(\theta) \tag{13}$$

Proposition 3. *Assuming that condition (10) is satisfied for all consumers with $\theta > 0$:*
(i) the social optimum entails higher participation than the private optimum: $N^c \geq N^{nc}$
(ii) the social optimum entails higher aggregate welfare than the private optimum: $W^c \geq W^{nc}$

proof: (i) Assume θ_0^c (resp. θ_0^{nc}) is the cutoff value of θ in the social (resp. private) optimum. Proposition (2iii) imposes the following inequality: $NPV^c(\theta_0^c) = 0 = NPV^{nc}(\theta_0^{nc}) \leq NPV^c(\theta_0^{nc})$. Since $NPV^j(\cdot)$ is increasing, $\theta_0^c \leq \theta_0^{nc}$. Hence, $N^c - N^{nc} = \int_{\theta_0^c}^{\theta_0^{nc}} \mathrm{d}F(\theta) \geq 0$. (ii) $W^c - W^{nc} = \int_{\theta_0^c}^{\theta_0^{nc}} NPV^c(\theta)\mathrm{d}F(\theta) + \int_{\theta_0^{nc}}^{+\infty}[U(\theta, s_\theta^c, q_\theta^c) - U(\theta, s_\theta^{nc}, q_\theta^{nc})]\mathrm{d}F(\theta) \geq 0$.

This is a very general formalization of the energy efficiency gap: In the presence of moral hazard, investments in energy efficiency entail too low a quality of installation and too few homeowners participate. This result holds under very general assumptions of perfect rationality and risk-neutrality. Concretely, the homeowner does not have the technical skills to judge whether the retrofit has been properly completed, although she is aware that any defects will deter the energy performance of the investment. Anticipating that the contractor is aware of her limitations, she will expect him to save on installation costs and perform the job poorly. Any claim that he will provide the highest quality, enabling her to maximize energy savings, will be considered ”cheap talk” by the homeowner. The contractor will not deviate from these expectations and indeed complete the lowest possible quality job. Quality will not be contractible and thus underprovided.

Appendix C discusses some comparative statics with respect to energy prices. This provides insights into the implementation of an instrument aimed at pricing energy-use externalities.

3. Policy solutions to energy efficiency moral hazards

In this section, we now examine some regulatory and incentive-based instruments that can be used to address moral hazard in energy efficiency markets.

3.1. Energy-savings insurance

Insurance is the most common way of addressing moral hazard problems. Energy-savings insurance or energy performance contracts typically have the contractor pay the consumer any shortfall in energy savings below a pre-agreed baseline. In our simple framework with no risk-aversion, the baseline is inessential and we can represent such contracts by merely having the contractor bear a share k of the energy bill:

$$U(\theta, s, q) \equiv \sum_{t=1}^{l} [\theta V(s) - (1-k)pE(s,q)]\,\delta^t - T + \epsilon \tag{14}$$

$$\Pi(q) \equiv T - C(q) - k\sum_{t=1}^{l} pE(s,q)\delta^t \tag{15}$$

According to Assumption (3), the tariff charged by the contractor is $T = C(q)+k\sum_{t=1}^{l} pE(s,q)\delta^t$, where $k\sum_{t=1}^{l} pE(s,q)\delta^t$ is the actuarially fair insurance premium.

A new, opposite principal-agent relationship superposes to the previous one: Since the contractor now provides insurance, he is a principal and the home owner-occupier is an agent. The implementation of this contract can be solved backward as a three-stage game played by the parties. In the third stage, each party determines non-cooperatively his or her own effort, given k and his or her belief about the other party's action. First-order conditions for maximization are:

$$\forall t \quad \theta V' \leq (1-k)pE_s \quad \text{with equality if} \quad s^i_\theta(k) > s_{min} \tag{16}$$

$$C' \geq -k\sum_{t=1}^{l} pE_q\delta^t \quad \text{with equality if} \quad q^i_\theta(k) > q_{min} \tag{17}$$

The optimal consumer's response is bounded above by a satiation value s_{max}[4]. By the Implicit Function Theorem, the insurance reaction functions $s^{**}_\theta(q,k)$ and $q^{**}(s,k)$ are both increasing in k:

$$\forall t \quad \frac{\mathrm{d}s^{**}_\theta}{\mathrm{d}k} = \frac{-pE_s}{\theta V'' - (1-k)pE_{ss}} > 0 \tag{18}$$

$$\frac{\mathrm{d}q^{**}}{\mathrm{d}k} = \frac{-\sum_{t=1}^{l} pE_q\delta^t}{C'' + k\sum_{t=1}^{l} pE_{qq}\delta^t} > 0 \tag{19}$$

[4]Satiation is needed in the model to handle full insurance ($k = 1$), which brings the marginal value of energy service in equation 16 to zero. It could be introduced as the argument of the maximum of a parabolic utility function. Alternatively, in our model, satiation is introduced as an upper bound on the value of s. This specification allows for more flexibility in the numerical section, without loss of generality.

The implementation of such a contract partly solves the moral hazard, as it induces the contractor to offer some quality (Equation (17)). At the same time, however, it gives rise to a second moral hazard: By lowering the homeowner's marginal value of energy service, it induces her to consume more energy. The energy service in Equation (16) is consumed to the socially optimal level defined by Equation (5) when the consumer is not insured ($k = 0$), whereas the quality in Equation (9) is offered to the socially optimal level defined by Equation (6) when the firm offers full insurance ($k = 1$). Since k cannot be simultaneously equal to 0 and 1, insurance cannot achieve the social optimum. At best, both parties will agree on an incomplete insurance contract $k \in (0, 1)$. We recover here the result established by Cooper and Ross (1985). For any insurance k, the agreement $(s^i_\theta(k), q^i_\theta(k))$ will be a Nash equilibrium determined by the intersection of each party's reaction function $s^{**}_\theta(q, k)$ and $q^{**}(s, k)$. These inputs will be higher than in the private optimum; however, their location relative to the social optimum is ambiguous.

Note that if consumer's types are imperfectly observable to the contractor, a screening issue arises: Consumers with the highest energy service use may self-select into the insurance contract that offers the highest energy savings coverage. Assuming this away, the optimal value $\hat{k}_\theta$ that sustains the Nash equilibrium to each type is determined cooperatively in the second stage of the game, so as to maximize joint expected surplus:

$$\forall\theta \quad \hat{k}_\theta = \underset{k\in[0,1]}{\arg\max}\ [U(\theta, s^i_\theta(k), q^i_\theta(k)) + \Pi(q^i_\theta(k))] \tag{20}$$

The first-order conditions for maximization in the second stage can be found in Appendix D. Lastly, in the first stage, the homeowner chooses whether or not to invest, depending on her net present value for the investment and given her beliefs about the contractor's action and the optimal insurance coverage.

Note that if the consumer were not optimizing her energy service and consuming a constant level of it (*e.g.*, a tenant who does not pay for her energy bill, or an employee in a commercial building), then the second moral hazard would not occur. The optimal insurance contract would feature full coverage and bring the parties to the social optimum.

3.2. **Minimum quality standard**

A social planner would like to get the contractor to provide the optimal level of quality and the homeowner to consume the optimal quantity of energy service. As the optimal level of quality is specific to each homeowner, this means that the social planner would have to monitor every contract, which would incur prohibitive administrative costs.

A minimum quality standard, which would translate into a perfectly enforced minimum labour requirement $\bar{q}$, is thus more likely to be implemented. Yet such an instrument may cause two classic types of deadweight loss. First, compliance with the standard still needs to be monitored, which generates increasing costs $M(\bar{q})$. These costs do not occur with energy-savings insurance, for the contract stipulates energy use, which is common knowledge. Second, minimum quality standards abstract from consumer heterogeneity. Take, for instance, an owner who visits her vacation home infrequently, thus consuming little heat there. The price of a high quality retrofit to save energy would be in excess of what is optimal to her. Now if the stringency of the standard is below what would be optimal to the consumer, as long as performance remains unobservable to her, the contractor will not offer more than the standard. Overall, $\bar{q}$ may be the optimal level of quality to one homeowner type, but it is suboptimal to all others (since, according to Proposition 1, the optimal quality is unique to each type θ). As a result, a uniform standard is strictly suboptimal over the whole population.

The best-implementable standard will be set at a value $\bar{q}$ that maximizes collective surplus, subject to the participation constraint:

$$\begin{aligned} &\underset{\bar{q}}{\text{Maximize}} && \left[\int_0^{\theta_0} U^0(\theta, s_\theta^0)\mathrm{d}F(\theta) + \int_{\theta_0}^{+\infty} \left[U(\theta, s_\theta^*(\bar{q}), \bar{q}) - M(\bar{q})\right]\mathrm{d}F(\theta)\right] \\ &\text{subject to} && NPV(\theta_0, s_{\theta_0}^*(\bar{q}), \bar{q}) - M(\bar{q}) \geq 0 \end{aligned} \tag{21}$$

As developed in Appendix E, the first-order condition for maximization will be:

$$\int_{\theta_0}^{+\infty} \left[\frac{\partial U(\theta, s_\theta^*(\bar{q}), \bar{q})}{\partial \bar{q}} - M'\right] \mathrm{d}F(\theta) = 0 \tag{22}$$

In words, the best-implementable standard will equalize the sum of marginal disutilities (net of marginal monitoring costs) of participants for whom the standard is too tight with the sum of marginal utilities (net of marginal monitoring costs) of participants who would have been willing to invest beyond the standard.

3.3. **Intervention rules with interacting energy-use externalities**

As we have just seen, addressing moral hazard problems through energy-savings insurance or quality standards can improve social welfare. Both instruments are, however, second-best. Uniform quality standards cannot eliminate the gap, because of the heterogeneity in consumer's valuation of energy services. Energy-savings insurance is incomplete because moral hazard is bilateral.

Yet public intervention to address moral hazard problems may not be systematically justified if moral hazard problems interact with energy-use externalities (*e.g.*, environmental externalities, energy security). Assume that every unit of energy consumed generates, over a time period of l years, an annual external cost $p_x > 0$ that is constant in present value. Expected consumer utility before and after investment is now:

$$\begin{cases} U_x^0(\theta, s) \equiv U^0(\theta, s) - p_x l E^0(s) \\ U_x(\theta, s, q) \equiv U(\theta, s, q) - p_x l E(s, q) \end{cases} \tag{23}$$

These new utility functions allow one to define new net present value NPV_x and aggregate welfare W_x functions as in equations (4) and (13), respectively. The optimal actions that internalize external costs are denoted by superscript x.

Proposition 4. *In a world subject to both energy-use externalities and moral hazard:*
(i) When energy-use externalities are internalized, it is desirable to also undo moral hazard problems: $W_x^{c,x} \geq W_x^{nc,x}$
(ii) If no consumer is prone to a genuine backfire rebound effect, then it is desirable to internalize energy-use externalities. This holds whether or not moral hazard problems are addressed: $\forall \theta \ E(s_\theta^c, q_\theta^c) \leq E^0(s_\theta^0) \Rightarrow W_x^{c,x} \geq W_x^c$ *and* $E(s_\theta^{nc}, q_\theta^{nc}) \leq E^0(s_\theta^0) \Rightarrow W_x^{nc,x} \geq W_x^{nc}$
(iii) If consumers are prone to neither a genuine nor a relative backfire rebound effect, then it is desirable to undo moral hazard problems. This holds even if energy-use externalities are not internalized: $\forall \theta \ E(s_\theta^c, q_\theta^c) \leq E(s_\theta^{nc}, q_\theta^{nc}) \leq E^0(s_\theta^0) \Rightarrow W_x^c \geq W_x^{nc}$

proof: See Appendix F.

As long as energy efficiency does not 'backfire', internalizing environmental externalities is desirable, regardless of whether or not the contracting parties overcome the moral hazard: Social welfare cannot be maximized if the parties do not account for the broader externalities produced by their actions. However, the reciprocal is not necessarily true: If environmental externalities are not (or cannot be) internalized, then it might be desirable to maintain, rather than undo, the moral hazard. This can actually occur if energy efficiency backfires. As a result, environmental externalities would be larger.

4. **A numerical illustration: Home weatherization**

The building sector is widely recognized as the largest and most cost-effective potential for energy savings and carbon dioxide emissions reduction (Levine *et al.*, 2007). Weatherization measures account for the bulk of this potential. According to McKinsey & Co. (2009), improvements of building shells and heating, ventilation and air conditioning (HVAC) systems could save 3 quadrillion end-use BTUs in the U.S. by 2020. Two-third of this amount would be achieved in existing homes. Yet this technical potential could remain partly untapped if the moral hazard problems attached to it are not addressed by government intervention.

4.1. **The sources of moral hazard in home weatherization**

Home weatherization technologies involve a significant installation input. If completed poorly by professionals, installation can be the source of many defects. This includes, for instance, an improper connection of ducts in HVAC systems, an imperfect filling of wall cavities before insulation or infiltrations around window panes. Detecting such defects is technically possible through a blower door test or thermographic screening. Yet these tests come at a substantial cost to the consumer[5]. Overall, there is little data available about the magnitude of these defects. Some analysis suggest it is sizeable on the extensive margin: As of 2008, only 15% of central air conditioning installations in existing dwellings met satisfactory quality specifications in California (Messenger, 2008). The intensive margin is particularly ill-documented. Metcalf and Hassett (1999) find that actual returns to attic insulation are around 10%, which is far from promises made by engineers and product manufacturers of 50%. The authors do not specifically investigate installation defects as an explanation for this gap, but their result provides suggestive evidence that poor quality is an issue in home weatherization.

The home energy retrofit industry is very fragmented. For instance, the HVAC industry in California is characterized by small firms offering low wages, with a very large number of quality problems reported (Zabin *et al.*, 2011). This can be interpreted as a ”bad” market

[5]Moreover, they are meant to be conducted before a retrofit, to help determine what measures should be undertaken. They are almost never conducted after the job is completed to check the quality of installation.

equilibrium involving low quality, similar to what we have described through the model as the private optimum. Moreover, Zabin *et al.*, (2011) also find that barriers to entry are low and that annual turnover is as high as 25%, which suggests that the competitive assumption made in our model is reasonable.

Various types of voluntary quality certifications exist in the marketplace, most notably those provided by the Building Performance Institute (BPI) and the Residential Energy Services Network (RESNET) in the U.S. These programs typically ensure that professional workers and contracting companies are trained to the best practices and that their performance is regularly tested. As of today, less than one percent of the professionals are certified, which suggests that the industry self-regulation has not been successful in addressing quality problems. In France, starting in 2014, public subsidies for home energy retrofits will be given only if the job is completed by a certified contractor. This "eco-conditionnality" rule is an interesting way of addressing jointly moral hazard and other problems[6].

Energy-savings insurance or energy performance contracts have been offered by energy service companies for about twenty years in the commercial sector (Mills, 2003). In contrast, these contracts are almost absent from the residential sector[7]. As we have seen through the model, such contracts may be welfare-improving but cannot achieve full efficiency due to the existence of a rebound effect, which typically ranges from 10 to 30% in space heating use (Sorrell *et al.*, 2009). In contrast, in the commercial sector, building occupants are expected to adopt a constant behavior, as they do not pay for the investment nor the operating costs. This lower uncertainty may explain why contractors are more likely to provide guarantee payments in the commercial sector.

This overview of the home retrofit industry shows that the model provides qualitative insights into some real-world facts. We now use it to conduct a quantitative assessment of the welfare implications of moral hazard problems. We focus on natural gas use for space heating in existing residential homes and investments in wall insulation in the U.S.

[6]The "other problems" justifying subsidies for energy efficiency can be either technology spillovers or energy-use externalities. In the latter case, though, subsidies are only a second-best solution (Giraudet and Quirion, 2008).

[7]One exception is the contracts offered by Green Homes America://www.greenhomesamerica.com/about-us/32-home-energy-audit-guarantee.aspx

4.2. **Functional forms used in the simulations**

The homeowner sets temperature s, measured in Fahrenheit (°F), above minimum comfort s_{min}. The value $V(\cdot)$ that she derives from this energy service is bounded above by V_{max}, which corresponds to a maximum budget dedicated to space heating. The function is increasing and concave, and takes the following form:

$$V(s) \equiv V_{max}\left(1 - \mathrm{e}^{-\alpha(s-s_{min})}\right) \quad \text{with} \quad s_{min} \leq s \leq s_{max}, \tag{24}$$

where $\alpha > 0$ is a calibrated parameter.

The use of natural gas E^0, measured in thousand cubic feet of natural gas (MCF), increases with indoor temperature (at an increasing rate) with a constant calibrated elasticity $\gamma > 1$:

$$E^0(s) \equiv \beta(s - s_{min})^{\gamma} \quad \text{with} \quad s_{min} \leq s \leq s_{max}, \tag{25}$$

where parameter $\beta > 0$ is calibrated so as to convert Fahrenheit degrees into thousand cubic feet.

Investment in wall insulation of efficiency $G(q)$ lowers energy use as follows:

$$E(s, q) \equiv (1 - G(q))\, E^0(s). \tag{26}$$

Efficiency is increasing in the quality q offered by the contractor (at a decreasing rate), within two limits $0 < G_{min} < G_{max} < 1$:

$$G(q) \equiv G_{min} + (G_{max} - G_{min})\left(1 - \mathrm{e}^{-\omega(q-q_{min})}\right) \quad \text{with} \quad q_{min} \leq q \leq q_{max}, \tag{27}$$

where $\omega > 0$ is a calibrated parameter.

The contractor bears a fixed cost K, which corresponds to a minimum labor input q_{min}. As the contractor provides the homeowner with a higher quality, he needs to have installers work longer and mobilize higher skills, which results in higher wages. As a result, cost increases quadratically in the number of worker.hours q:

$$C(q) \equiv K + \rho(q - q_{min}) + \frac{\phi}{2}\,(q - q_{min})^2 \quad \text{with} \quad q_{min} \leq q \leq q_{max}, \tag{28}$$

where $\rho > 0$ and $\phi > 0$ are calibrated parameters.

4.3. **Data and calibration**

Homeowners' characteristics are drawn from the U.S. Department of Energy's Residential Energy Consumption Survey (RECS) of 2009. We focus on information on indoor temperature, energy use, energy expenditure and income contained in the online database. We first extract a preliminary sample of 4,306 U.S. households who own and occupy their house (variable KNOWRENT=1) and pay for natural gas for space heating (variable PGASHEAT=1). We then remove households who declare a winter daytime temperature below 60°F or above 80°F (variable TEMPHOME) and thereby obtain a working sample of 4,266 households. This sample covers 35% of the complete dataset. Summary statistics are provided in Table 1.

To isolate moral hazard problems from other market or behavioral failures, we assume that homeowners discount future energy expenditures at a "failure-free" rate of 7%, over the full physical lifetime of an insulation project (35 years). The environmental damages caused by carbon dioxide emissions associated with natural gas use are valued at $33/t$CO_2$.

Fewer data are available to parameterize the supply side of the insulation market. We therefore use best guesses based on anecdotal evidence found in the grey literature and drawn from discussions with practitioners. Our assumptions are detailed in Table 2.

In the RECS sample, the annual fraction of homeowners investing in insulation is 3.4%[8]. Our model is calibrated so that this rate is replicated in the private optimum and can be doubled at best. That is, participation among potential investors is set to 50% in the private optimum. A participation of 100% in the model can thus be interpreted as an annual insulation rate of 6.8% in the total population. With this calibration, we find non-energy benefits of insulation of $2,035. All calibration targets are outlined in Table 3 and the calibration procedure is detailed in Appendix G.

The temperature distribution found in the RECS sample is best fitted with a log-normal distribution of homeowners' types with parameters $\mu = 0$ and $\sigma = 1$. Yet we do not intend to reproduce such a large heterogeneity, which may be partly driven by variables omitted in our model. Therefore, in the simulations, we assume a narrower distribution of homeowners' types with log-normal parameters $\mu = 0$ and $\sigma = 0.25$, from the 0.5th percentile ($\theta = 0.53$) to the 99.5th percentile ($\theta = 1.90$). The model fit is illustrated in Figure 1.

[8] 6.8% of the population declare having insulation installed in the last two years (variable AGEINS=1). Note that the 3.4% rate is close to 2.9%, which would be the annual rate if investment occurred once every 35 years.

With these structural and numerical assumptions, the homeowner of type $\theta = 1$ is both the median of the distribution and the marginal participant in the private optimum. The model satisfies the necessary and sufficient conditions for existence and uniqueness of private and social equilibria (Proposition 1).

4.4. **Quantification of the energy efficiency gap**

Simulation results are given in Table 4 and illustrated in the figures. In Figure 2, various equilibria are mapped *à la* Jaffe and Stavins[9] (1994), so as to visualize the trade-offs between economic efficiency and energy efficiency. Without internalization of energy-use externalities, the private optimum generates modest improvements in either welfare or energy efficiency, compared to the equilibrium before investment. In contrast, when the contractor cooperates to undo the moral hazard, both welfare and energy efficiency improvements become substantial: Average energy efficiency moves from 3% to 27% and lifetime discounted welfare with environmental damages increases by \$1,723 (or \$1,249 if environmental damages are not accounted for). This number is several times larger than the cost of a home energy audit, estimated on average at \$347 (Palmer *et al.*, 2013). Therefore, government intervention intended to undo the moral hazard could be beneficial to society.

Further improvements along both the energy efficiency and welfare dimensions can occur if environmental damages are internalized through a carbon price. In Jaffe and Stavins' words, the social optimum is then moved from the "Narrow economists' optimum" to the "True social optimum". The welfare gains from undoing the moral hazard (\$1,723) are one order of magnitude larger than those from internalizing energy-use externalities (\$162). This proportion reflects the difference between the marginal inefficiency due to moral hazard, namely the unit of energy that could have been cut by optimal investment (valued at energy price $p = \$11.14/\text{MCF}$), and the social cost of environmental damages (valued at $p_{CO_2} = \$33/\text{tCO}_2 = \$1.54/\text{MCF}$).

Note that the higher welfare level in the true social optimum is not general to the model, but due to the absence of backfire rebound effects in our calibration (see Proposition 4iii). Indeed, as reported in Table 4, the "genuine" rebound effect is 31% in the private optimum and 33% in the social optimum. In addition, we find a "relative" rebound effect of 33%, meaning that 33% of energy efficiency gains are taken back when the economy moves from

[9] The authors have refined their conceptual diagram over the years. The version we specifically refer to first appeared in Jaffe *et al.* (2004).

the private to the social optimum. These numbers are in the top of the 10-30% range reported by Sorrell *et al.* (2009) for space heating rebound effects.

The moral hazard market failure can be restated as an implicit discount rate of 20% on average over the population. This is done by solving and averaging the discount rates that, for each θ, match the quality enjoyed in the social optimum with the net present value enjoyed in the private optimum, discounted at 7% by assumption. The value obtained is in the middle of the 10-32% range reported for investments affecting the thermal integrity of dwellings (Train, 1985).

Figure 2 illustrates, with the median homeowner, how equilibria are formed through reaction function intersections. The reaction functions are mildly upward slopping. The consumer's energy service varies by no more than 2 Degrees Fahrenheit with the quality of installation. Such a sensitivity is consistent with the values found in the literature (Hirst *et al.*, 1985). While the quality offered by the contractor is flat a minimum if he behaves non-cooperatively, it mildly increases with consumer's energy service if he behaves cooperatively. As predicted by Proposition 2, the social optimum implements both a higher quality and a higher energy service than the private optimum. The figure also pictures some comparative statics of the energy price. Pricing energy-use externalities shifts the consumer's reaction function inward and the contractor's upward (*c.f.* Equations 32 and 33). As discussed in Appendix C, the final location of the equilibrium - at a higher quality and a lower energy service - is not general but specific to model parameters.

4.5. **Efficiency of policy instruments**

To capture the ability of minimum quality standards and energy-savings insurance to specifically address moral hazard problems, we examine here welfare effects in a world without environmental damages.

Were homeowner types perfectly observable, the government would implement standards corresponding to each homeowner's optimal quality. Moreover, insuring contractors could design optimal contracts for every homeowner (Figure 6). On average, such contracts stipulate a coverage of 33%. As illustrated in Figure 4, insurance shift the reaction functions toward higher parties' actions (*c.f.* Equations 18 and 19). For the median homeowner, the resulting equilibrium entails a quality level that is intermediate between the social and private one and an energy service that is higher than in the social optimum. Again, this positioning is contingent upon our calibration, not general to the model. Under full insurance, the contractor offers the socially optimal quality but the consumer faces zero marginal energy expenditures and thus sets her energy service at corner s_{max}.

In practice, homeowners' types are unobservable and screening issues arise: The insuring contractor cannot offer each homeowner her best contract, just like the government cannot implement as many standards as there are homeowner types. Rather, both the contractor and the government seek the uniform instrument that best fits aggregate homeowners' tastes. The best implementable instrument can be found analytically, as exemplified by Equation 21 for the uniform standard. Yet numerical resolution of such an equation is tedious, so policy implementation is more likely to result from some *tâtonnement* around policy parameters. Such a process is depicted as parametric curves in Figure 7.

Let us assume first that monitoring a minimum quality standard is costless. Increasing the standard from 24 to 47 worker.hours increases both economic efficiency and energy efficiency. Further increasing the stringency increases energy efficiency but not economic efficiency: The standard is getting too tight for most of the people. The value of the best-implementable standard is very close to the quality that is optimal to the median homeowner (47.1 worker.hours). As shown in Figure 5, such a standard is too tight to the 5th percentile homeowner (type $\theta = .66$), who would have been best-off with a standard of 45.6 worker.hours. It is too loose to the 95th percentile homeowner (type $\theta = 1.51$), who would have been best-off with a standard of 48.4 hours. Therefore, there is a narrow quality range in which the standard can be set so that average welfare is very close to the social optimum. The fact that screening issues induce only negligible welfare losses is due to the curvatures of reaction functions being very mild.

A similar pattern is observed with uniform insurance. Increasing insurance coverage up to 30% increases both energy efficiency and economic efficiency. Between 30% and 40%, the best-implementable contract is very close to the situation where consumers are all offered their best contract. Again, welfare losses due to screening are negligible. Energy efficiency then increases at the expense of economic efficiency up to a coverage of 60%, a situation described by Jaffe and Stavins as a "Technologist's optimum". Beyond that point, energy efficiency starts decreasing too, until the curve crashes, for a coverage of $k = 100\%$, at the welfare level enjoyed before investment. That is, under full insurance, the total cost paid to the contractor is so high that no homeowner can be left with a positive net present value, hence none invests. Otherwise, indoor temperature would be set at corner $s_{max} =$ 80°F, leading to an annual natural gas use of 79 MCF. This would imply a prohibitive lifetime discounted energy bill of $11,442, fully borne by the contractor and passed on to the homeowner as an insurance premium.

Overall, we find that the average welfare enjoyed with the best-implementable standard is $290 higher than with best-implementable insurance (and $720 higher if environmental damages are accounted for). Unlike insurance, a standard entails monitoring costs, estimated for each realization at $347 from Palmer *et al.* (2013). In practice, only a fraction of realizations can be randomly monitored, so this estimate is an upper bound of the true monitoring cost. When accounting for monitoring costs, the welfare difference (without environmental damages) between the standard and insurance becomes ambiguous. However, our assumption of insurance contracts running over 35 years is probably too sanguine and overestimates the welfare gains from insurance. Therefore, the superiority of minimum quality standards over energy-savings insurance seems to be a robust result.

5. Discussion

The model was built upon standard assumptions, in order to carefully isolate the sources and consequences of moral hazard problems in energy efficiency decisions. In this section, we reexamine each of the behavioral and organizational assumptions made to question the generality of the model.

Considering risk-aversion would be a natural extension of the model. As a matter of fact, the performance of energy efficiency technologies depends on volatile factors, such as weather conditions or energy prices, to which consumers are likely to have adverse reactions. Moreover, the home retrofit industry is made up of small industries that have little room to diversify risks (Lutzenhiser, 1994). In our model, risk-averse homeowners would expect higher energy expenditures than certainty equivalent, hence demand less energy service. Risk-averse firms would respond with a lower quality in the social optimum. Overall, the introduction of risk-aversion on both sides of the market would reduce the size of the energy efficiency gap compared to a riskless situation.

The hidden actions examined in this analysis may propagate as hidden information in subsequent principal-agent relationships. That is, the contractor's failure can be internalized as the homeowner's by prospective home buyers or renters, leading to non-capitalization of energy efficiency performance in home sale prices or rental contracts. One way to introduce non-capitalization in the model is to assume that consumers compute the net present value of energy efficiency investments over their expected home occupancy duration (typically 10 years) rather than the conventional lifetime of the product (assumed to be 35 years in our simulations). With such an assumption, we find that in a world with environmental damages, the average welfare at the private optimum would be \$11,848, instead of \$24,270 if energy efficiency is fully capitalized. The welfare difference (\$12,422) is the amount by which our estimate of the welfare gains from undoing moral hazard (\$1,723) would increase were the non-capitalization problem addressed at the same time. The average implicit discount rate under both moral hazard and the non-capitalization problem would be 26%, instead of 20% without the latter. Altogether, the welfare losses from moral hazard between investors and contractors may be strongly amplified if they propagate in subsequent sale transactions.

Our model focused on imperfect information rather than imperfect rationality. As is increasingly discussed in the literature, however, consumers may value energy savings in a way that is inconsistent with perfect rationality (Allcott and Greenstone, 2012; Gillingham and Palmer, 2014). Introducing energy savings undervaluation biases in our model would, again, widen the energy efficiency gap: A cooperative contractor would internalize consumer's undervaluation and supply even less quality than in the perfect rationality benchmark.

Our market structure assumption too could be relaxed[10]. Even though the industry seems to satisfy the condition of perfect competition with free entry, one might expect firms to exert monopoly power locally. Home energy retrofits are very specific to a bundle of home and homeowner characteristics, hence do not lend themselves to arbitrage. Hence, a monopolist could perfectly price discriminate, all the more if screening issues are unimportant, as is the case in the model. Perfect discrimination would not change equilibrium quantities in the model, but only the surplus repartition: In the social optimum, the retrofit price charged to the median consumer would be \$2,764 by a competitive firm and \$4,060 by a perfectly discriminating monopolist (*c.f.* Table 4).

Instead of assuming homogeneous firms that all fail to offer quality in equilibrium, we could assume heterogeneous firms. Within our informational structure, a fringe of firms may adopt reputation or signaling strategies and supply a better, or even optimal quality. Such private forces would lead to reduce the size of the energy efficiency gap.

The refinements discussed above change our assessment of the energy efficiency gap in different directions. Yet altogether, they would not dramatically change the key insights of the model. We therefore conduct a naive extrapolation of our simulations to the U.S. population. Recall that the homeowners using natural gas for space heating covered 35%

[10]Here we only refer to the structure of the market for energy efficiency. The effect of market power in the energy market is given by the energy price comparative statics discussed in the paper.

of the RECS dataset. Moreover, we confined our attention to 6.8% of that subpopulation (hence 2.4% of the complete dataset), the fraction of households we considered capable to undertake insulation each year. Applying these shares to a number of U.S. households of 115 million (U.S. Census Bureau value for 2008-2012), our analysis covered approximately 2.8 million households. If all enjoy, on average, 9 MCF of annual energy savings and $1,723 of lifetime discounted welfare gains, then the total benefits of undoing moral hazard could be 25 billion cubic feet of annual natural gas savings and $5 billion of welfare gains.

6. **Conclusion**

In this paper, we provide a framework to think about how moral hazard problems can generate an energy efficiency gap, and how this market failure may interact with energy-use externalities. Taking home energy retrofits as an example, we show that if the quality of installation offered by a retrofit contractor is unobserved to the homeowner, the contractor will cut quality in equilibrium. This leads to a suboptimal level of energy efficiency along both the intensive and extensive margins: The quality offered to consumers is too low and there are too few consumers investing.

Numerical simulations show that moral hazard is consistent with implicit discount rates in the 20-30% range. It also induces an energy efficiency gap that is potentially larger than the one induced by energy-use externalities. This insight is relatively new, as most of the literature has concluded that it was hard to find market failure explanations for the abnormally high implicit discount rates observed in energy efficiency decisions (Gillingham *et al.*, 2009; Allcott and Greenstone, 2012). While most studies have focused on the role of possible undervaluation of energy efficiency by consumers, ours underlines the importance of considering the behavior of the firms.

Our analysis provides motivation for policies that would go beyond the internalization of energy-use externalities. This recommendation holds as long as consumers are not prone to a ”backfire” rebound effect - a reasonable hypothesis (Sorrell, 2009; Gillingham *et al.*, 2013; Borenstein, 2014). When addressing the moral hazard, the first-best outcome can only be attained to the extent that energy performance and consumer preferences can be made perfectly observable. Since no technology can meet that goal at an affordable cost yet, government intervention will only generate second-best outcomes: Minimum quality standards do not address consumer heterogeneity and energy-savings insurance raise a second moral hazard. However, our numerical results suggest that the former can bring social welfare very close to its optimal level. Similarly, even with a modest coverage, insurance products can deliver welfare gains that should not be disregarded.

7. Appendices

7.1. Appendix A: Comparative statics with respect to consumer type θ

Applying the Implicit Function Theorem to equation (5):

$$\forall t \quad \frac{\mathrm{d}s^*_\theta}{\mathrm{d}\theta} = \frac{-V'}{\theta V'' - pE_{ss}} > 0 \tag{29}$$

Therefore, for any given quality q offered by the contractor, a higher valuation of energy service shifts the consumer's reaction function upward:

$$\forall q,\ \forall \theta_1 > \theta_2 \quad s^*_{\theta_1}(q) > s^*_{\theta_2}(q) \tag{30}$$

As long as condition (10) is satisfied, new equilibria are determined with the properties below:

Proposition 6. *If condition (10) is satisfied for two participating consumers of types θ_1 and θ_2, with $\theta_1 > \theta_2$, then the higher θ implies higher actions by either contracting party, in either equilibrium:*
(i) $q^{nc}_{\theta_1} = q^{nc}_{\theta_2} = q_{min}$
(ii) $s^{nc}_{\theta_1} \geq s^{nc}_{\theta_2}$
(iii) $s^c_{\theta_1} \geq s^c_{\theta_2}$
(iv) $q^c_{\theta_1} \geq q^c_{\theta_2}$.

proof: (i) is straightforward. (ii) Combined with (30), it implies: $s^{nc}_{\theta_1} = s^*_{\theta_1}(q^{nc}_{\theta_1}) \geq s^*_{\theta_2}(q^{nc}_{\theta_2}) = s^{nc}_{\theta_2}$. (iii) Likewise, (30) implies, for all s, $s^*_{\theta_1}(q^*(s)) \geq s^*_{\theta_2}(q^*(s))$. In particular, $s^c_{\theta_1} = s^*_{\theta_1}(q^*(s^c_{\theta_1})) \geq s^*_{\theta_2}(q^*(s^c_{\theta_1}))$. From (10), $s^*_{\theta_2}(q^*(\cdot))$ is increasing with slope lower than 1. Any point that is greater than its image by $s^*_{\theta_2}(q^*(\cdot))$ is thus greater than the fixed point of $s^*_{\theta_2}(q^*(\cdot))$: $\forall a > s^c_{\theta_2},\ s^*_{\theta_2}(q^*(a)) - s^*_{\theta_2}(q^*(s^c_{\theta_2})) < a - s^c_{\theta_2} \Leftrightarrow s^*_{\theta_2}(q^*(a)) < a$. Therefore, $s^c_{\theta_1} \geq s^c_{\theta_2}$. (iv) Lastly, since $g^*(\cdot)$ is increasing, $q^c_{\theta_1} = g^*(s^c_{\theta_1}) \geq g^*(s^c_{\theta_2}) = q^c_{\theta_2}$.

7.2. Appendix B: Participation to investment

For any equilibrium situation j, participation will depend on the limits of the net present value function, the sign of which is indeterminate:

$$NPV^j(\theta) \equiv \sum_{t=1}^{l} \underbrace{\left[\theta\left(V(s^j_\theta) - V(s^0_\theta)\right) - p\left(E(s^j_\theta, q^j_\theta) - E^0(s^0_\theta)\right)\right]}_{\geq 0} \delta^t \underbrace{-T + \epsilon}_{\leq 0} \tag{31}$$

The right inequality is given by Assumption 1v. The left inequality comes from the following inequalities: $\theta V(s^0_\theta) - pE^0(s^0_\theta) \leq \theta V(s^0_\theta) - pE(s^0_\theta, q^j_\theta) \leq \theta V(s^j_\theta) - pE(s^j_\theta, q^j_\theta)$. The former is due to technological assumptions about E and E^0 and the latter is due to s^j_θ maximizing U.

Thanks to proposition 6, equilibrium actions s^j_θ and q^j_θ decrease with θ. As they are bounded below by s_{min} and q_{min}, the limit of $NPV(\theta)$ when θ tends toward zero is finite.

if $\lim\limits_{\theta\to 0} NPV(\theta) \geq 0$ then all consumers participate. Participation is given by $N^j \equiv \int_0^{+\infty} \mathrm{d}F(\theta) = 1$.

If $\lim\limits_{\theta\to 0} NPV(\theta) < 0$ and $\lim\limits_{\theta\to +\infty} NPV(\theta) > 0$ then by equation 11, there exists a unique cutoff type θ_0, as discussed in the text.

If $\lim\limits_{\theta\to 0} NPV(\theta) < 0$ and $\lim\limits_{\theta\to +\infty} NPV(\theta) \leq 0$ then participation is nil. In this case, the gross utility gains accruing to the homeowner never offset the increase in the tariff charged by the contractor.

7.3. **Appendix C: Energy price comparative statics**

Applying the Implicit Function Theorem to equations 5 and 6, we see that an increase in energy price shifts reaction functions $s^*_\theta(\cdot)$ downward and $q^*(\cdot)$ upward:

$$\forall t \quad \frac{\mathrm{d}s^*_\theta}{\mathrm{d}p} = \frac{E_s}{\theta V'' - pE_{ss}} < 0 \tag{32}$$

$$\frac{\mathrm{d}q^*}{\mathrm{d}p} = \frac{-\sum\limits_{t=1}^{l} pE_q\delta^t}{C'' + \sum\limits_{t=1}^{l} pE_{qq}\delta^t} > 0 \tag{33}$$

By the same reasoning as in proposition 6, a higher energy price entails a higher energy service in private equilibrium. But optimal actions cannot be compared unambiguously in the private and social equilibria.

The influence of p on NPV^*, established by the Envelope Theorem, depends on the consumer's reaction to higher energy efficiency:

$$\frac{\mathrm{d}NPV^*}{\mathrm{d}p} = -\sum_{t=1}^{l} \left[E(s^*_\theta, q^*_\theta) - E^0(s^0_\theta) \right] \delta^t \tag{34}$$

As long as energy efficiency investments decrease energy use for all consumers, the net present value is increasing in energy price. By the same type of reasoning as in Proposition 6, this leads to a higher participation and a higher average welfare. This conclusion is reversed if all consumers are subject to a genuine backfire rebound effect, i.e., $\forall\theta\ E(s^*_\theta, q^*_\theta) > E^0(s^0_\theta)$. In this case, higher energy prices decrease participation and average welfare.

7.4. **Appendix D: Optimal insurance coverage**

The first-order condition for finding the optimal insurance contract from equation (20) is:

$$\frac{\mathrm{d}s^{**}_\theta}{\mathrm{d}k} \sum_{t=1}^{l} \left[\theta V' - pE_s\right] \delta^t - \frac{\mathrm{d}q^{**}}{\mathrm{d}k} \left[C' + \sum_{t=1}^{l} pE_q \delta^t \right] = 0 \tag{35}$$

Plugging in equations (16) and (17) and further simplifying gives the equation that solves the optimal coverage $\hat{k}$:

$$\forall t \quad kE_s \frac{\mathrm{d}s^{**}_\theta}{\mathrm{d}k} + (1-k)E_q \frac{\mathrm{d}q^{**}}{\mathrm{d}k} = 0 \tag{36}$$

7.5. **Appendix E: Optimal minimum quality standard**

Assuming that the cutoff type exists and is unique, the constraint in the optimization program (21) is binding. The program can be solved by simply maximizing the objective function and assuming that θ_0 is an implicit function $\theta_0(\bar{q})$ defined by the constraint. Applying the Leibniz integral rule and the Envelope Theorem leads to the following first-order condition for maximization:

(37)
$$\frac{\mathrm{d}\theta_0}{\mathrm{d}\bar{q}}\Big(U^0(\theta_0(\bar{q}), s^0_\theta) - U(\theta_0(\bar{q}), s^*_\theta(\bar{q}), \bar{q}) + M(\bar{q})\Big) + \int_{\theta_0(\bar{q})}^{+\infty}\left[\frac{\partial U(\theta, s^*_\theta(\bar{q}), \bar{q})}{\partial \bar{q}} - M'\right]\mathrm{d}F(\theta) = 0$$

Recognizing that $U^0(\theta_0(\bar{q}), s^0_\theta) - U(\theta_0(\bar{q}), s^*_\theta(\bar{q}), \bar{q}) = -NPV(\theta_0(\bar{q}), s^*_{\theta_0}(\bar{q}), \bar{q})$ and using the binding constraint leads to the result (equation 22).

Note that if participation to investment is nil without the standard, no standard will be welfare-improving. In contrast, if participation is full without the standard, the constraint will not be binding and the optimal standard will be defined by the following first-order condition:

(38)
$$\int_0^{+\infty}\left[\frac{\partial U(\theta, s^*_\theta(\bar{q}), \bar{q})}{\partial \bar{q}} - M'\right]\mathrm{d}F(\theta) = 0$$

7.6. **Appendix F: Proof of Proposition 4**

(i) For all θ, since $(s^{c,x}_\theta, q^{c,x}_\theta)$ maximizes U_x in the social setting, $U_x(\theta, s^{c,x}_\theta, q^{c,x}_\theta) \geq U_x(\theta, s, q)$ for all (s, q), and for $(s^{nc,x}_\theta, q^{nc,x}_\theta)$ in particular. Likewise, we have $U^0_x(\theta, s^{0,x}_\theta) \geq U^0_x(\theta, s^0_\theta)$. By Proposition (4), it follows that $W^{c,x}_x \geq W^{nc,x}_x$.

(ii) Again, for all θ, since $(s^{c,x}_\theta, q^{c,x}_\theta)$ maximizes U_x in the social setting, $U_x(\theta, s^{c,x}_\theta, q^{c,x}_\theta) \geq U_x(\theta, s^c_\theta, q^c_\theta)$. In addition, we have $NPV_x(\theta) = NPV(\theta) - p_x L[E(s,q) - E^0(s)]$. Assume θ^x_0 is the cutoff type in an equilibrium where both energy-use externalities and moral hazard are addressed, while θ_0 is the cut-off type in an equilibrium where only moral hazard problems are addressed. We have $NPV_x(\theta^x_0) = 0 = NPV(\theta_0)$. In the absence of a genuine backfire rebound effect, we thus have $NPV_x(\theta^x_0) = 0 \leq NPV_x(\theta_0)$. Since NPV is increasing in θ, $\theta^x_0 \leq \theta_0$, that is, participation is higher if externalities are internalized. The difference in aggregate welfare between the two equilibria is $\Delta W = \int_0^{\theta^x_0} \Delta U^0_x \mathrm{d}F(\theta) + \int_{\theta^x_0}^{\theta_0}[U_x(\theta, s^x_\theta, q^x_\theta) - U^0_x(\theta, s^0)]\mathrm{d}F(\theta) + \int_{\theta_0}^{+\infty} \Delta U_x \mathrm{d}F(\theta)$. The first and third integrands of the right-hand side are positive (see proof (i) just above). The second integrand is also positive, since $\forall\theta \geq \theta^x_0\ U_x(\theta, s^x_\theta, q^x_\theta) \geq U^0_x(\theta, s^{0x}) \geq U^0_x(\theta, s^0)$. Therefore, aggregate welfare is larger when externalities are internalized: $W^{c,x}_x \geq W^c_x$. The exact same reasoning leads to $W^{nc,x}_x \geq W^{nc}_x$. This is because since $(s^{nc,x}_\theta, q^{nc,x}_\theta)$ maximizes U_x in the private setting, $U_x(\theta, s^{nc,x}_\theta, q^{nc,x}_\theta)$ is greater than $U_x(\theta, s, q)$ for any other actions s and q determined in a private setting, e.g., $(s_\theta^{nc}, q_\theta^{nc})$.

(iii) Assume θ_0^c (resp. θ_c^{nc}) is the cutoff type in the social (resp. private) optimum. From proposition (4i), we have $\theta_0^c \leq \theta_0^{nc}$. Therefore, the aggregate welfare difference between the two situations is $\Delta W_x = \int_{\theta_0^c}^{\theta_0^{nc}} NPV_x(\theta, s_\theta^c, q_\theta^c)\mathrm{d}F(\theta) + \int_{\theta_0^{nc}}^{+\infty}[U_x(\theta, s_\theta^c, q_\theta^c) - U_x(\theta, s_\theta^{nc}, q_\theta^{nc})]\mathrm{d}F(\theta)$. In the absence of a genuine backfire rebound effect, the first term of the right-hand side is positive (see proof (ii) just above). In the absence of a relative backfire rebound effect, the second term of the right-hand side is also positive. To see this, note that $\forall\theta\ E(s_\theta^c, q_\theta^c) \leq E(s_\theta^{nc}, q_\theta^{nc}) \Rightarrow -p_x LE(s_\theta^c, q_\theta^c) \geq -p_x LE(s_\theta^{nc}, q_\theta^{nc})$. This, added to $U(\theta, s_\theta^c, q_\theta^c) \geq U(\theta, s_\theta^{nc}, q_\theta^{nc})$ (which is given by definition of the maximum) leads to $U_x(\theta, s_\theta^c, q_\theta^c) \geq U_x(\theta, s_\theta^{nc}, q_\theta^{nc})$. To conclude, the aggregate welfare difference is positive: $W_x^c \geq W_x^{nc}$.

7.7. **Appendix G: Model calibration**

Parameters α, β and γ are computed so as to allow the model to replicate calibration targets 1, 2 and 3. For $\theta = 1$ and $s^0 = 69°$F, this leads to:

(39)
$$\begin{cases} V'(s^0) - pE_s^0(s^0) = 0 \\ E^0(s^0) = 50 \\ \frac{\mathrm{d}E^0\left(s^0_{\theta=1}(p)\right)}{\mathrm{d}p}\frac{p}{E^0\left(s^0_{\theta=1}(p)\right)} = -0.4 \end{cases} \Leftrightarrow \begin{cases} V_{max}\alpha\exp\left(-\alpha(s^0 - s_{min})\right) - p\gamma\beta(s^0 - s_{min})^{\gamma-1} = 0 \\ \beta(s^0 - s_{min})^\gamma = 50 \\ \frac{\gamma}{1-\gamma-\frac{V_{max}\alpha^2\exp\left(-\alpha(s^0-s_{min})\right)}{p\gamma\beta(s^0-s_{min})^2}} = -0.4 \end{cases}$$
$$\Leftrightarrow \begin{cases} \alpha = 0.28 \\ \beta = 5.32 \\ \gamma = 1.02 \end{cases}$$

Parameter ϵ, representing the non-energy benefits of insulation, is computed so as to allow the model to replicate calibration target 4. As the marginal investing homeowner has type $\theta = 1$, ϵ is such that $NPV(\theta = 1, s_{\theta=1}^{nc}, q_{\theta=1}^{nc}) = 0$, which leads to $\epsilon = \$2,035$.

Parameter ω is computed so as to allow the model to replicate calibration target 5:

(40) $$G(q_{max}) = .99G_{max} \Leftrightarrow G_{min} + (G_{max} - G_{min})(1 - \omega(q_{max} - q_{min}) = .99G_{max} \Leftrightarrow \omega = .09$$

Parameters ρ and ϕ are computed so as to allow the model to replicate calibration targets 6 and 7:

$$\begin{cases} C'(q_{min}) = 15 \\ C'(q_{max}) = 30 \end{cases} \Leftrightarrow \begin{cases} \rho = 15 \\ \rho + \phi(q_{max} - q_{min}) = 30 \end{cases} \Leftrightarrow \begin{cases} \rho = 15 \\ \phi = .16 \end{cases} \tag{41}$$

8. Acknowledgements

We thank, without involving, Tiger Adolf, Rick Diamond, Matt Golden, Evan Mills and Carol Zabin for discussions that inspired this work. We are grateful to Frédéric Branger, Larry Goulder, Franck Lecocq, Marie-Laure Nauleau, Philippe Quirion, Jim Sweeney and Adrien Vogt-Schilb for their comments on earlier results. We also thank seminar audiences at CIRED, Stanford University, Université Paris X-Nanterre, Paris School of Economics, as well as participants to the WEAI Conference in San Francisco, the AEAA Conference in Seattle, the USAEE Conference in Austin, the AERE Summer Conference in Banff and the EAERE Conference in Toulouse for their feedback.

9. References

Allcott, H., M. Greenstone, 2012, ”Is There an Energy Efficiency Gap?”, *Journal of Economic Perspectives*, 26(1):3-28

Allcott, H., Mullainathan, S., Taubinsky, D., 2014. ”Energy policy with externalities and internalities.” *Journal of Public Economics* 112, 72-88.

Aroonruengsawat, A., Auffhammer, M., Sanstad, A.H., 2012. ”The Impact of State Level Building Codes on Residential Electricity Consumption.” *The Energy Journal* 33, 31-52.

Boomhower, J., Davis, L.W., 2014. ”A credible approach for measuring inframarginal participation in energy efficiency programs.” *Journal of Public Economics* 113, 67-79.

Borenstein, S., 2014, ”A Microeconomic Framework for Evaluating Energy Efficiency Rebound And Some Implications”, *The Energy Journal*, forthcoming.

Cooper, R., T.W. Ross, 1985, ”Product Warranties and Double Moral Hazard”, *The RAND Journal of Economics*, 16(1):103-113

Gillingham, K., R.G. Newell, K. Palmer, 2009, ”Energy Efficiency Economics and Policy”, *Annual Review of Resource Economics*, 1(1):597-620

Gillingham, K., M. Harding, D. Rapson, 2012, ”Split Incentives in Residential Energy Consumption”, *The Energy Journal*, 33(2):37-62

Gillingham, K., M. J. Kotchen, D. S. Rapson, G. Wagner, 2013, ”Energy policy: The rebound effect is overplayed”, *Nature*, 493(7433):475-476

Gillingham, K., K. Palmer, 2014, ”Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Evidence”, *Review of Environmental Economics and Policy*, 8(1):18-38

Giraudet, L.-G., P. Quirion, 2008, ”Efficiency and distributional impacts of tradable white certificates compared to taxes, subsidies and regulations”, *Revue d'économie politique*, 118(6):885-914

Hirst, E., D.White, R. Goeltz, 1985, ”Indoor temperature changes in retrofit homes”, *Energy*, 10(7):861-870

Houde, S., 2014, ”How Consumers Respond to Environmental Certification and the Value of Energy Information”, NBER Working Paper No. 20019

Howarth, R. B., B. Andersson, 1993, ”Market barriers to energy efficiency”, *Energy Economics*, 15(4):262-272

Huntington, H.G., A.H. Sanstad, L.J. Schipper, 1994, ”Editors' introduction”, pp. 795-797 *in* Huntington, Sanstad and Schipper (eds), ”Markets for Energy Efficiency, Special Issue,” *Energy Policy*, 22(10)

Ito, K., 2013, ”Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program”, NBER Working Paper No. 19485

Jacobsen, G.D., Kotchen, M.J., 2011. ”Are Building Codes Effective at Saving Energy? Evidence from Residential Billing Data in Florida.” *Review of Economics and Statistics* 95, 34-49.

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

Jaffe, A. B., R. G. Newell, R. N. Stavins, 2004, ”Economics of energy efficiency”, *Encyclopedia of Energy*, 2:79-90

Jessoe, K., Rapson, D., 2014. ”Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use”. *American Economic Review* 104, 1417-1438.

Levine, M., D. Ürge-Vorsatz, K. Blok, L. Geng, D. Harvey, S. Lang, G. Levermore, A. Mongameli Mehlwana, S. Mirasgedis, A. Novikova, J. Rilling, H. Yoshino, 2007, *2007: Residential and commercial buildings*, *in* Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Lutzenhiser, L., 1994, ”Innovation and organizational networks. Barriers to energy efficiency in the US housing industry”, *Energy Policy*, 22(10):867-876

McKinsey & Co., 2009, *Unlocking Energy Efficiency in the U.S. Economy*

Messenger, M., 2008, *Strategic Plan to Reduce the Energy Impact of Air Conditioners*, California Energy Commission Staff Report (CEC-400-2008-010)

Metcalf, G. E., K. A. Hassett, 1999, ”Measuring the energy savings from home improvement investments: Evidence from monthly billing data”, *Review of economics and statistics*, 81(3):516-528

Mills, E., 2003, ”Risk transfer via energy-savings insurance”, *Energy Policy*, 31(3):273-281

OMB, [U. S. Office of Management and Budget], n.d. Informing regulatory decisions: report to Congress on the costs and benefits of federal regulations and unfunded mandates on state, local, and tribal entities. (No. Appendix D, OMB Circular A-4).

Palmer, K., M., Walls, H. Gordon, T. Gerarden, 2013, ”Assessing the energy-efficiency information gap: results from a survey of home energy auditors”, *Energy Efficiency*, 6(2): 271-292

Sorrell, S., 2004, ”Understanding barriers to energy efficiency”, pp.25-93 *in* Sorrell S., E. O’Malley, J. Schleich, S. Scott (eds), *The Economics Of Energy Efficiency: Barriers to Cost-Effective Investment*, Cheltenham: Edward Elgar

Sorrell, S., J. Dimitropoulos, M. Sommerville, 2009, ”Empirical estimates of the direct rebound effect: A review”, *Energy Policy*, 37(4):1356-1371

Sorrell, S., 2009. ”Jevons’ Paradox revisited: The evidence for backfire from improved energy efficiency.” *Energy Policy* 37, 1456-1469.

Train, K., 1985, ”Discount rates in consumer’s energy-related decisions: a review of the literature”, *Energy*, 10(12):1243-1253

White House, 2013, *Technical Support Document: - Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866*

Zabin, C., T.W. Lester, J. Halpern-Finnerty, ”Chapter Four: Case studies of the HVAC, residential retrofit, and commercial lighting controls sectors”, p.89-118, *in* Donald Vial Center on Employment in the Green Economy, *California Workforce Education and Training Needs Assessment for Energy Efficiency, Distributed Generation, and Demand Response*, 2011, University of California, Berkeley, Institute for Research on Labor and Employment

Variable	Unit	RECS Entry	Median	Mean	Standard Deviation	Minimum	Maximum
Temperature when someone is home during the day (winter)	°F	TEMPHOME	69.0	69.3	3.4	60.0	80.0
Natural Gas cost for space heating, 2009	$	DOLNGSPH	562	624	386	33	3,591
Natural Gas usage for space heating, 2009	MCF	$\frac{\text{BTUNGSPH}}{1.023*10^{-3}}$	50.2	55.0	34.0	0.6	337.2
Price paid for natural gas for space heating	$/MCF	$\frac{\text{DOLNGSPH}*1.023}{\text{BTUNGSPH}*10^{3}}$	11.14	11.95	5.69	3.75	190.73
Gross household income, 2009	$	MONEYPY	65,000	78,928	50,727	2,500	170,000
Income share dedicated to natural gas for space heating		$\frac{\text{DOLNGSPH}}{\text{MONEYPY}}$	0.82%	1.54%	3.20%	0.02%	65.92%

TABLE 1. **Summary Statistics of the RECS Sample ($n = 4,266$).** 'Natural gas cost for space heating', given in the dataset in thousand BTU (variable BTUNGSPH), is converted here in MCF. 'Gross household income' is given in the dataset as 24 income ranges (variable MONEYPY); we identify each income range with its upper value and assume an average income of $170,000 for the top category, which is consistent with U.S. Census Bureau 2009 data for owner-occupiers.

Parameter	Symbol	Value	Unit	Source
Minimum indoor temperature	s_{min}	60	°F	0.8th percentile of the RECS preliminary sample
Maximum indoor temperature	s_{max}	80	°F	99.9th percentile of the RECS preliminary sample
Minimum labour input	q_{min}	24	worker.hours	One workday = 24 worker.hours, *e.g.*, three installers working 8 hours a day (Best guess)
Maximum labour input	q_{max}	72	worker.hours	Three workdays (Best guess)
Maximum valuation of energy service	V_{max}	2,816	\$	The 95th percentile of the income share dedicated to space heating in the RECS sample is 4.3%. Applying this fraction to the median income of the sample (\$65,000) leads to a maximum budget for space heating of \$2,816.
Minimum energy efficiency of insulation	G_{min}	5%		(Best guess)
Maximum energy efficiency of insulation	G_{max}	30%		(Best guess)
Fixed cost of wall insulation	K	2,400	\$	Unit cost of \$1/sf applied to a representative dwelling of 2,400 square feet (Best guess)
Physical lifetime of insulation investment	l	35	years	(Best guess)
Discount rate	$r = \frac{1-\delta}{\delta}$	7%		Value recommended to assess private investment (U.S. OMB).
Price of natural gas	p	11.14	\$/MCF	Median price of the RECS sample
Carbon price	p_{CO_2}	1.54	\$/MCF	Equivalent to a social cost of carbon of \$33/tCO_2 in 2010, which is the value recommended for impact analysis in the U.S. (White House, 2013)

TABLE 2. **Model parameters**

	Calibration target	Expression	Target value	Source
1	Optimal temperature before investment to the median homeowner	$s^0_{\theta=1}$	69°F	Median temperature of the RECS sample
2	Optimal annual energy use before investment to the median homeowner	$E^0(s^0_{\theta=1})$	50 MCF	Median annual natural gas use for space heating of the RECS sample
3	Price-elasticity of energy demand before investment to the median homeowner	See App.G	-0.4	Middle value of the [-0.03;-0.76] range found in the literature by Gillingham *et al.* (2009) for short-term price-elasticities of natural gas use
4	Participation to insulation investment in the private optimum	See App.G	50%	See text (Section 4.3).
5	Expected energy savings at maximum quality level	$G(q_{max})$	$99\% G_{max}$	
6	Wage for insulation workers at minimum quality level	$C'(q_{min})$	\$15/hour	According to the U.S. Bureau of Labour and Statistics, the median pay for insulation workers was \$16.88/hour in 2010. According to Zabin *et al.* (2011), the lower range of insulation wages in California is 10-15\$/hour
7	Wage for insulation workers at maximum quality level	$C'(q_{max})$	\$30/hour	Upper range of the values reported by Zabin *et al.* (2011) for California

TABLE 3. **Model calibration targets.** Parameters α, β, γ, ω, ρ, ϕ and ϵ are calibrated so as to allow the model to replicate these targets. The procedure is detailed in Appendix G.

	Without carbon price				With carbon price			
Model outputs, averaged over the population	Before investment $j=0$	Social optimum $j=c$	Private optimum $j=nc$	Optimal insurance $j=i$	Before investment $j=0,x$	Social optimum $j=c,x$	Private optimum $j=nc,x$	Formula
Welfare without environmental damages ($)	27,134	28,397	27,148	28,112	26,945	28,254	26,968	W^j
Welfare with environmental damages ($)	24,206	25,992	24,270	25,260	24,419	26,154	24,486	W^j_x
Annual natural gas use for space heating (MCF)	49.5	40.7	48.7	48.2	42.7	35.5	41.9	$E^j \equiv \int_0^{\theta_0^j} E^0(s^0_\theta)\mathrm{d}F(\theta) + \int_{\theta_0^j}^{+\infty} E(s^j_\theta, q^j_\theta)\mathrm{d}F(\theta)$
Homeowners' equilibrium temperature (°F)	68.3	69.4	68.5	70.6	67.1	68.2	67.3	$\int_0^{\theta_0^j} s^0_\theta \mathrm{d}F(\theta) + \int_{\theta_0^j}^{+\infty} s^j_\theta \mathrm{d}F(\theta)$
Energy efficiency		26.9%	2.5%	22.9%		27.4%	4.7%	$G^j \equiv \int_{\theta_0^j}^{+\infty} G(q^j_\theta)\mathrm{d}F(\theta)$
Rebound effect		33%	31%	89%		38%	61%	$1-(1-E^j/E^0)/G^j$
Cutoff type of the marginal participant		0.18	1.00	0.26		0.19	0.68	θ_0^j
Participation rate		99.5%	49.0%	99.5%		99.5%	93.6%	N^j
Contractor's equilibrium quality (wr.hr)		47.1	24.0	37.8		24.0	24.0	$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} q^j_\theta \mathrm{d}F(\theta)$
Zero profit price of insulation ($)		2,829	2,400	2,637		2,875	2,400	$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} C(q^j_\theta)\mathrm{d}F(\theta)$
Homeowner's net present value ($)		1,269	28	983		1,743	85	$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} NPV^j(\theta)\mathrm{d}F(\theta)$
Wage for insulation workers ($/wr/hr)		22.21	15.00	19.32		22.84	15.00	$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} C'(q^j_\theta)\mathrm{d}F(\theta)$
Insurance premium ($)				2,314				$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} \hat{k}_\theta \sum_{t=1}^{l} pE(s^j_\theta, q^j_\theta)\delta^t \mathrm{d}F(\theta)$
Insurance optimal coverage				33%				$\frac{1}{N^j}\int_{\theta_0^j}^{+\infty} \hat{k}_\theta \mathrm{d}F(\theta)$

TABLE 4. **Simulation results, averaged over the population of total mass 1.** Unit 'wr.hr' is 'worker.hour'. 'Energy efficiency' is averaged over the whole population. The average over participants is obtained by dividing 'Energy Efficiency' by 'Participation rate'. Adding 'Zero profit price of insulation' and 'Homeowner's net present value' gives the price that would be charged by a perfectly discriminating monopolist.

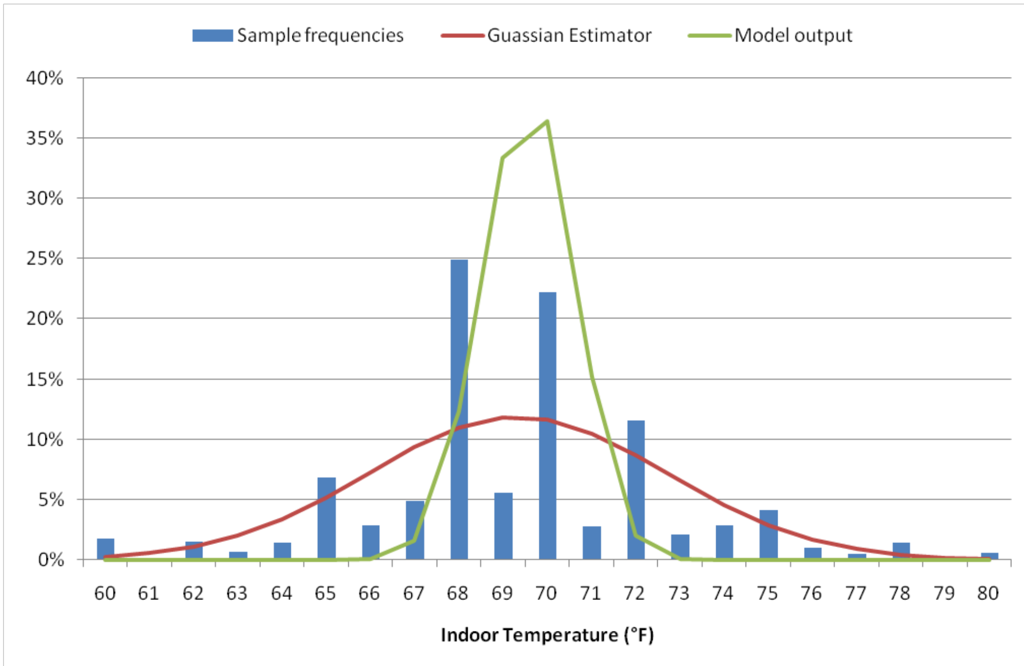

FIGURE 1. **Model fit to RECS sample data.** The Gaussian estimator is the normal distribution of temperature with parameters $\mu = 69.3$ and $\sigma = 3.4$, the mean and standard deviation of the RECS sample (Table 1). The model output is the probability distribution function of s^0_θ, calculated with the triangle method and assuming a log-normal distribution of θ with parameters $\mu = 0$ and $\sigma = 0.25$.

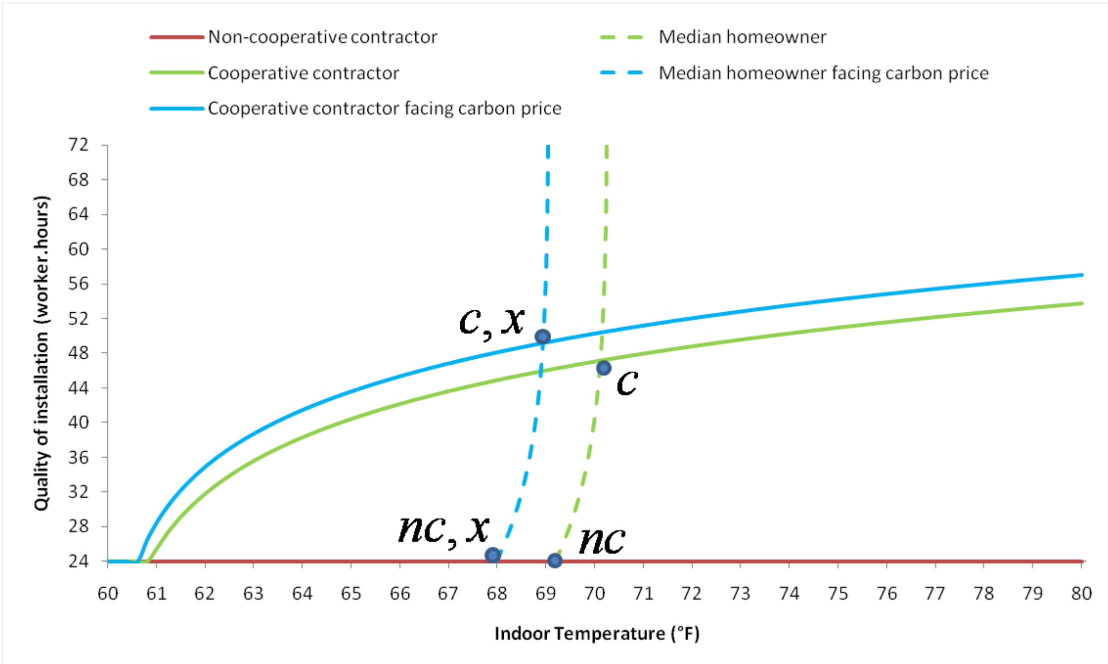

FIGURE 2. **Reaction functions, with a homeowner of median type $\theta = 1$.** (c) refers to the social optimum, (nc) to the private optimum and (c, x) and (nc, x) to the same optima in the presence of a carbon price of \$33/tCO_2.

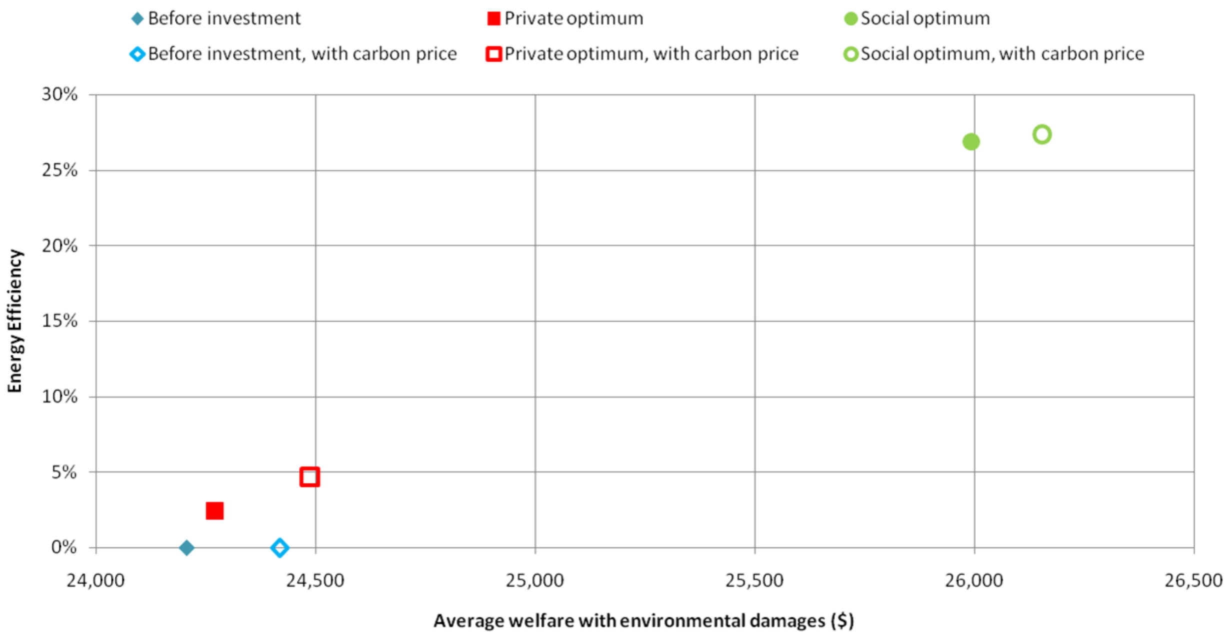

Figure 3. **The energy efficiency gap.** The horizontal axis represents lifetime discounted average welfare over the population. Environmental damages are valued at $33/tCO$_2$. The vertical axis represents average energy efficiency over the population (with the value 0% attributed to non-participating homeowners).

.

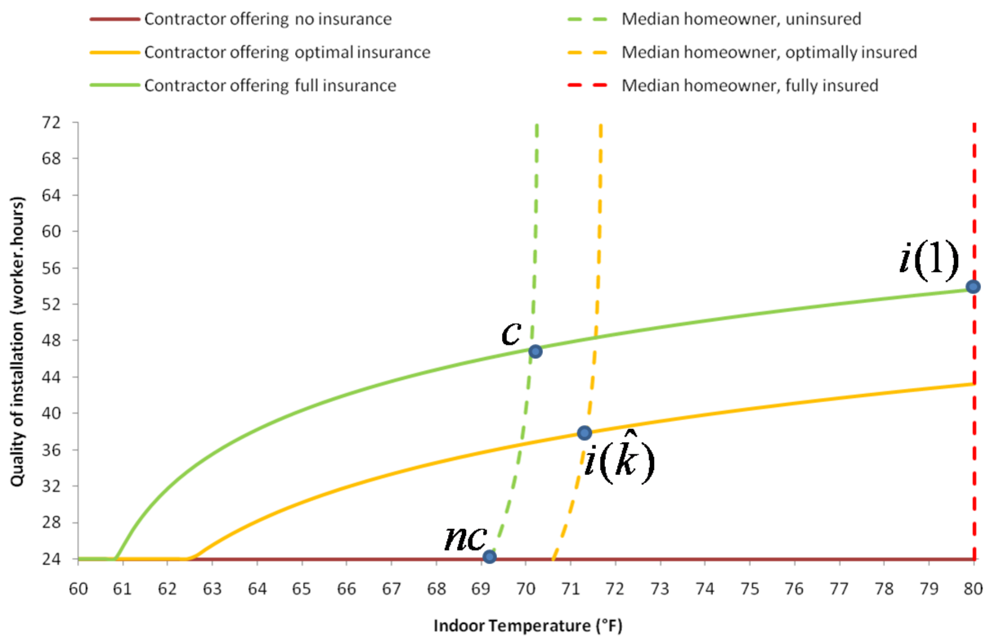

Figure 4. **Reaction functions under energy-savings insurance, with a homeowner of median type $\theta = 1$.** (c) refers to the social optimum, (nc) to the private optimum, $(i(\hat{k}))$ to the equilibrium induced by insurance with optimal coverage $\hat{k}$ and $(i(1))$ to the full insurance equilibrium.

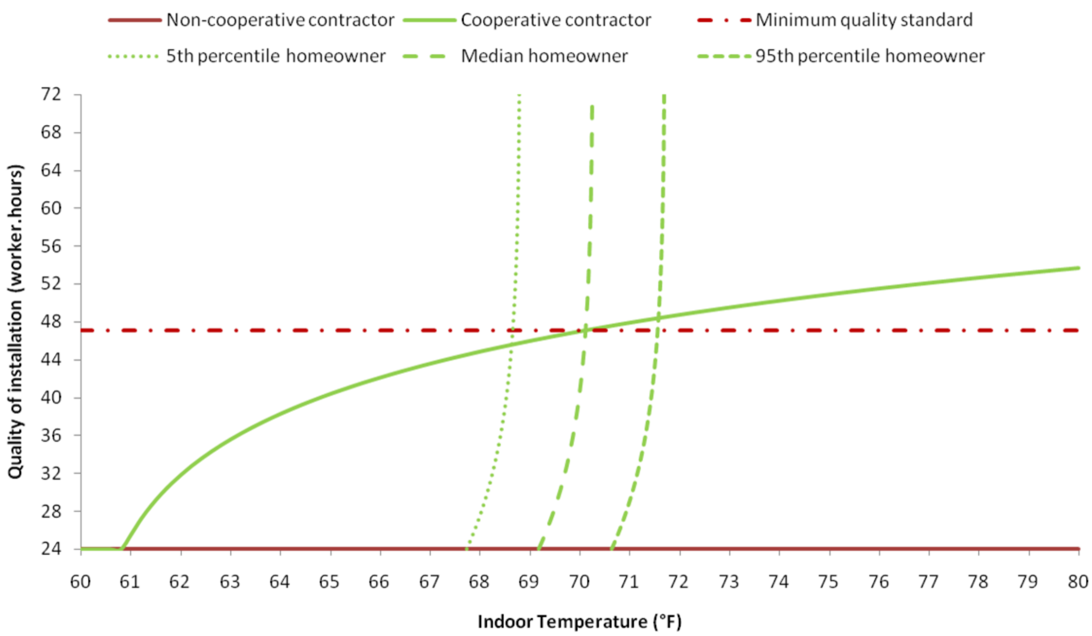

FIGURE 5. **Reaction functions with a minimum quality standard.** The standard is optimal to the median homeowner ($\theta = 1$), but suboptimal to all others. For instance, it is too tight to the 5th percentile of the homeowners' distribution ($\theta = 0.66$) and too loose to the 95th percentile ($\theta = 1.51$).

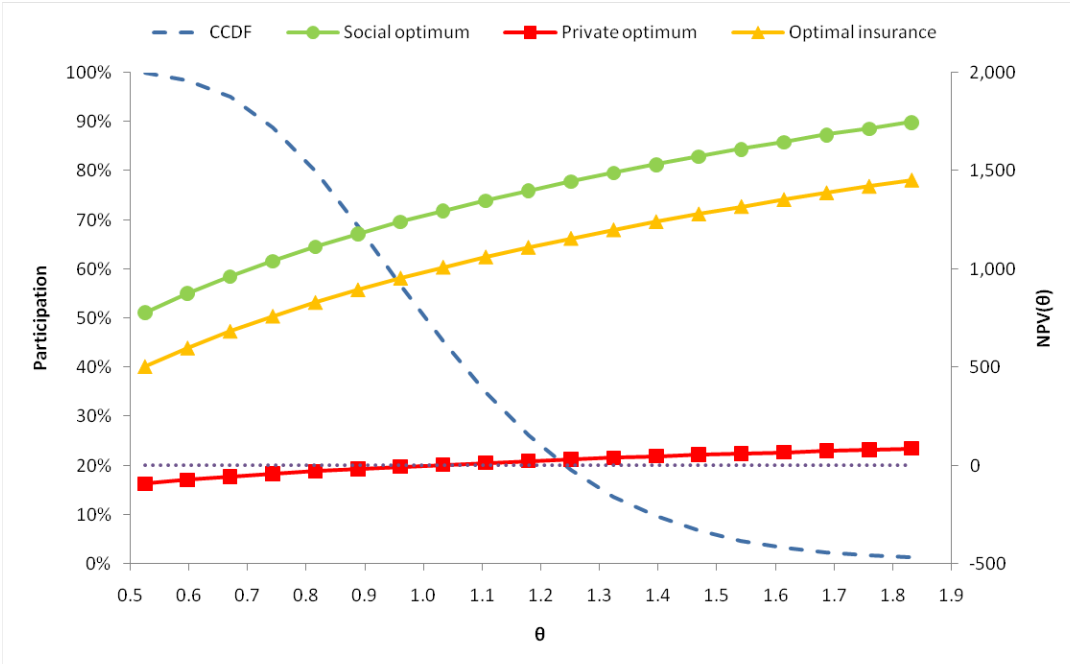

FIGURE 6. **Net present value and participation with respect to homeowner's type θ.** The net present value (without environmental damages) of investment in insulation reads on the right vertical axis. The intersection of each curve with the zero horizontal axis determines the cutoff type θ_0 of the marginal participant in investment. For each cutoff type on the horizontal axis (from the 5th to 95th percentile of the θ distribution), participation across the population is determined by the value of the complementary cumulative distribution (CCDF) of θ, which reads on the left vertical axis.

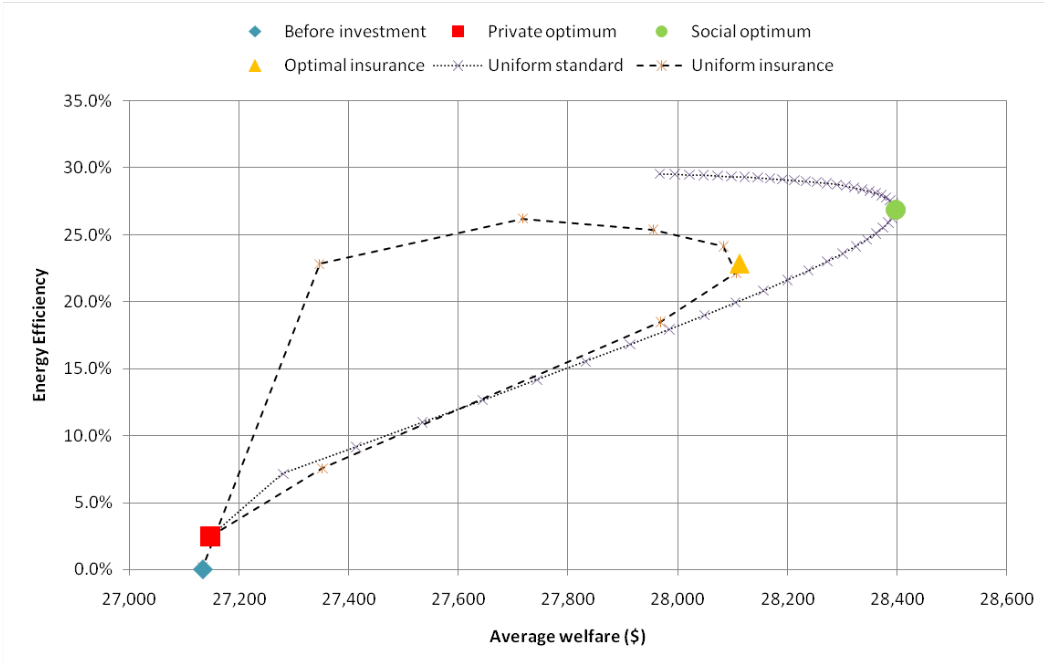

FIGURE 7. **Economic and energy efficiency of policy instruments.** The welfare considered here does not take environmental damages into account. Each mark of the uniform standard parametric curve represents an additional worker.hour of labour requirement, from q_{min} to q_{max}. The stringency of the standard increases counter-clockwise. Each mark of the uniform insurance curve represents an incremental 10% of insurance coverage, from 0 to 100%. Coverage increases counter-clockwise.