

Moral Hazard and the Energy Efficiency Gap: Theory and Evidence

Louis-Gaëtan Giraudet,^{*} Sébastien Houde,[†] Joseph Maher[‡]

Online Appendix

December 9, 2016

CONTENTS

A1. Setup	2
A2. Supply-side moral hazard	4
A2.1. Energy efficiency contract	4
A2.2. Energy efficiency gap at the market level	6
A2.3. Comparative statics	8
A3. Policy solutions	9
A3.1. Energy-savings contracts and double moral hazard	9
A3.2. Minimum quality standard	12
A3.3. Interaction between moral hazard and energy-use externalities	13
A4. Sufficient statistics	15
A4.1. Deadweight loss from supply-side moral hazard	15
A4.2. Marginal welfare effect of a quality standard	16
A4.3. Marginal welfare effect of energy-savings contracts	16

^{*}Centre International de Recherche sur l'Environnement et le Développement (CIRED), Ecole des Ponts ParisTech. Email: giraudet@centre-cired.fr.

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

[‡]National Socio-Environmental Synthesis Center (SESYNC) University of Maryland. Email: jmaher@sesync.org.

This online appendix provides formal assumptions, additional propositions and complete proofs to the theoretical part developed in Section 2 of the main text. It also provides background for the sufficient-statistic formula discussed in Section 4. Graphical illustrations are produced with the model specifications detailed in Section 4.

A1. Setup

The homeowner maximizes lifetime-discounted utility $U^0(\cdot, \cdot)$ with respect to an intertemporal energy-service vector \mathbf{s} :

$$(1) \quad U^0(\theta, \mathbf{s}) \equiv \sum_{t=1}^l [\theta V_t(s_t) - p_t E_t^0(s_t)] (1+r)^{-t}$$

where $V(\cdot)$ is the value derived from energy service (e.g., heating comfort), $\theta > 0$ is a taste parameter (e.g., a cold-intolerant person having a high θ), $E^0(\cdot)$ is energy use, p is the price of energy, l is investment lifetime and r is the discount rate.

The homeowner's lifetime-discounted net utility after retrofit $U(\theta, \mathbf{s}, q)$ reads

$$(2) \quad U(\theta, \mathbf{s}, q) \equiv \sum_{t=1}^l [\theta V_t(s_t) - p_t E_t(s_t, q)] (1+r)^{-t} - T + \epsilon$$

where q is the quality of installation offered by the contractor, $T > 0$ is the price of the retrofit and ϵ is the value derived from non-energy attributes (e.g., acoustic or aesthetic benefits if positive, inconvenience due to job completion if negative).

In what follows, we assume time invariance of energy price, technology and homeowner value function. We remove t subscripts and consider vector \mathbf{s} as a scalar s constant over time. We further simplify notation with a discount factor Γ such that:

$$(3) \quad \Gamma \equiv \Gamma(r, l) \equiv \sum_{t=1}^l (1+r)^{-t} = \frac{1 - (1+r)^{-l}}{r}$$

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:

$$(4) \quad \Pi(q) \equiv T - C(q)$$

The following assumptions hold (subscripts denote partial derivatives):

Assumption 1. Technology

(i) At constant homeowner behavior s , investment reduces energy use: $E(s, q) < E^0(s)$ $\forall q \geq q_{min}$, where q_{min} is the minimum input.

(ii) Contracting parties' actions have opposite effects: $E_s^0 > 0$, $E_s > 0$ and $E_q < 0$

(iii) Energy savings exhibit decreasing returns: $-E_{ss}^0 \leq 0$, $-E_{ss} \leq 0$ and $-E_{qq} \leq 0$

(iv) Contracting parties' actions are substitutes: $E_{qs} < 0$ and $E_s < E_s^0$

(v) Non-energy benefits are not sufficient to motivate investment: $\epsilon \leq C(q_{min})$

Assumption 2. Preferences

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 structure

The industry is competitive with free entry: $\Pi(q) = 0$.

Corollary. $T = C(q)$

Remark. Assumptions 1(i)-(v) are mild: The energy service has a convex effect on energy use, and quality has diminishing returns on energy savings. Moreover, both factors impede each other: The marginal increase in energy savings due to increased quality is larger when the 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 energy use due to increased energy service is lower when the quality installed is high rather than low.

Assumptions 2(i)-(iii) are meant to be as standard as possible, in order to isolate the moral-hazard problem from possibly interacting market failures and behavioral anomalies.

Assumption 3 is not essential. Whatever the structure of the market, home energy retrofits are very specific to a bundle of home and homeowner characteristics, and hence do not lend themselves to arbitrage. A monopolist could thus perfectly price discriminate. This would not change equilibrium quantities in the model, but only the surplus repartition.

The energy efficiency contract can be modeled as a two-stage game, of which the homeowner is the principal and the contractor is the agent. In the first stage of the game, a homeowner of type θ invests if the j -equilibrium net present value of investment $NPV^j(\theta)$ is positive, given her beliefs about her future optimal energy service s_θ^j and the optimal quality q_θ^j she will be offered by the contractor:

$$(5) \quad NPV^j(\theta) \equiv U(\theta, s_\theta^j, q_\theta^j) - U_0(\theta, s_\theta^0) \geq 0$$

In the second stage, both agents determine their own action given their belief about the other party's action.

A2. Supply-side moral hazard

A2.1. Energy efficiency contract

We compare the equilibrium of the game under two information structures, namely perfect and asymmetric information. The resulting equilibria can be seen as the social and private optimum, respectively.

Perfect information (PI). The contract between the contractor and a homeowner of type θ is set cooperatively so as to maximize joint surplus, subject to boundary conditions $s \geq s_{min}$ and $q \geq q_{min}$. The optimal actions s_{θ}^{PI} and q_{θ}^{PI} that solve the first-order conditions for maximization⁴ below will be such that their marginal benefit (in terms of value to the homeowner and cost savings to the firm) equates their marginal effect on homeowner's energy bill:

$$(6) \quad \forall t \quad \theta V' \begin{cases} = pE_s & \text{if } s_{\theta}^{PI} > s_{min} \\ < pE_s & \text{otherwise} \end{cases}$$

$$(7) \quad C' \begin{cases} = -pE_q & \text{if } q_{\theta}^{PI} > q_{min} \\ < -pE_q & \text{otherwise} \end{cases}$$

The perfect-information equilibrium $(s_{\theta}^{PI}, q_{\theta}^{PI})$ 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:

$$(8) \quad \forall t \quad \frac{ds_{\theta}^*}{dq} = \frac{pE_{qs}}{\theta V'' - pE_{ss}} > 0$$

$$(9) \quad \frac{dq^*}{ds} = \frac{-pE_{sq}}{C''/\Gamma + pE_{qq}} > 0$$

⁴Throughout the objective functions are well-behaved and the first-order conditions discussed are necessary and sufficient conditions for maximization.

Asymmetric information (AI). The agreement is no longer cooperative. Both parties maximize their private value, given their beliefs about the 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 homeowner, this does not hold for the contractor. He fails to internalize the benefits his action delivers to the homeowner and simply chooses the level of quality q^{AI} that minimizes his cost:

$$(10) \quad \forall s \quad q^{AI}(s) = \arg \min_{q \geq q_{min}} C(q) = q_{min}$$

Lemma 1. *For a participating homeowner of given type θ :*

- (i) *the asymmetric-information equilibrium $(s_\theta^{AI}, q_\theta^{AI})$ exists and is unique*
- (ii) *the perfect-information equilibrium $(s_\theta^{PI}, q_\theta^{PI})$ exists and is unique if and only if:*

$$(11) \quad \frac{dq}{ds_\theta^*} > \frac{dq^*}{ds}$$

Proof. (i) The asymmetric-information equilibrium is uniquely defined as $(s_\theta^*(q_{min}), q_{min})$. (ii) Likewise, if for at least one agent his or her optimal cooperative action is a corner solution, then the perfect-information equilibrium is uniquely defined. If optimal actions are interior for both agents, condition (11) 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. \square

Figure A1 illustrates the reaction function equilibria. The following proposition states that the perfect and asymmetric-information equilibria will involve unambiguous locations:

Proposition 1. *Assuming condition (11) holds, a participating homeowner of given type θ :*

- (i) *is offered less quality under asymmetric information: $q_\theta^{AI} \leq q_\theta^{PI}$*
- (ii) *uses less energy service under asymmetric information: $s_\theta^0 < s_\theta^{AI} \leq s_\theta^{PI}$*
- (iii) *faces a lower net present value under asymmetric information: $NPV^{AI}(\theta) \leq NPV^{PI}(\theta)$*

Proof. (i) For a given θ , $q_\theta^{PI} \geq q_{min} = q_\theta^{AI}$. (ii) Since $s_\theta^*(\cdot)$ is increasing, $s_\theta^{PI} = s_\theta^*(q_\theta^{PI}) \geq s_\theta^*(q_\theta^{AI}) = s_\theta^{AI}$. For all s , $E_s^0 > E_s$ implies $U_s > U_s^0$. Therefore, assuming interior solutions: $U_s^0|_{s_\theta^0} = 0 = U_s|_{s_\theta^{AI}} > U_s^0|_{s_\theta^{AI}}$. Since U^0 is concave in s , U_s^0 is decreasing in s and $s_\theta^{AI} > s_\theta^0$. (iii) Comparing net present values $NPV^{PI}(\cdot)$ and $NPV^{AI}(\cdot)$ is equivalent to comparing the utility functions after investment $U(\theta, s_\theta^{PI}, q_\theta^{PI})$ and $U(\theta, s_\theta^{AI}, q_\theta^{AI})$. Under the assumption of perfect competition, the utility after investment is equivalent to the joint surplus. Therefore, the net present value of investment is maximized under perfect information: $NPV^{PI}(\theta) \geq NPV^{AI}(\theta)$. \square

Recall that $E(s_\theta^{PI}, q_\theta^{PI})$, $E(s_\theta^{AI}, q_\theta^{AI})$ and $E(s_\theta^0)$ 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. *Investment-induced backfire*

An investment-induced backfire rebound effect occurs if energy use after investment is larger than before investment: $s > s^0$ and $E(s, q) > E(s^0)$.

Definition 2. *Policy-induced backfire*

A policy-induced backfire rebound effect occurs between two policy 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 $E(s^H, q^H) > E(s^L, q^L)$.

A2.2. Energy efficiency gap at the market level

We now turn to a continuum of homeowners of mass 1. Applying the Implicit Function Theorem to Equation 6:

$$(12) \quad \forall t \quad \frac{ds_\theta^*}{d\theta} = \frac{-V'}{\theta V'' - pE_{ss}} > 0$$

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

$$(13) \quad \forall q, \forall \theta_h > \theta_l \quad s_{\theta_h}^*(q) > s_{\theta_l}^*(q)$$

This effect is illustrated in Figure A2.

As long as condition (11) is satisfied, new equilibria are determined with the properties below:

Lemma 2. *If condition (11) is satisfied for two participating homeowners of types θ_h and θ_l , with $\theta_h > \theta_l$, then the higher θ implies higher actions by either contracting party, in either equilibrium:*

- (i) $q_{\theta_h}^{AI} = q_{\theta_l}^{AI} = q_{min}$
- (ii) $s_{\theta_h}^{AI} \geq s_{\theta_l}^{AI}$
- (iii) $s_{\theta_h}^{PI} \geq s_{\theta_l}^{PI}$
- (iv) $q_{\theta_h}^{PI} \geq q_{\theta_l}^{PI}$.

Proof. (i) is straightforward. (ii) Combined with (13), it implies: $s_{\theta_h}^{AI} = s_{\theta_h}^*(q_{\theta_h}^{AI}) \geq s_{\theta_l}^*(q_{\theta_l}^{AI}) = s_{\theta_l}^{AI}$. (iii) Likewise, (13) implies, for all s , $s_{\theta_h}^*(q^*(s)) \geq s_{\theta_l}^*(q^*(s))$. In particular, $s_{\theta_h}^{PI} = s_{\theta_h}^*(q^*(s_{\theta_h}^{PI})) \geq s_{\theta_l}^*(q^*(s_{\theta_h}^{PI}))$. From (11), $s_{\theta_l}^*(q^*(\cdot))$ is increasing with slope lower than 1. Any point that is greater than its image by $s_{\theta_l}^*(q^*(\cdot))$ is thus greater than the fixed point of $s_{\theta_l}^*(q^*(\cdot))$: $\forall a > s_{\theta_l}^{PI}$, $s_{\theta_l}^*(q^*(a)) - s_{\theta_l}^*(q^*(s_{\theta_l}^{PI})) < a - s_{\theta_l}^{PI} \Leftrightarrow s_{\theta_l}^*(q^*(a)) < a$. Therefore, $s_{\theta_h}^{PI} \geq s_{\theta_l}^{PI}$. (iv) Lastly, since $g^*(\cdot)$ is increasing, $q_{\theta_h}^{PI} = g^*(s_{\theta_h}^{PI}) \geq g^*(s_{\theta_l}^{PI}) = q_{\theta_l}^{PI}$. \square

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

$$(14) \quad \frac{dNPV^j}{d\theta} = [V(s_{\theta}^j) - V(s_{\theta}^0)] \Gamma$$

As $V(\cdot)$ is increasing and $\forall \theta$ $s_{\theta}^j > s_{\theta}^0$, the net present value of investment strictly increases with θ . This is illustrated in Figure A3. For any equilibrium situation j , participation will depend on the limits of the net present value function, the sign of which is indeterminate:

$$(15) \quad NPV^j(\theta) \equiv \underbrace{[\theta (V(s_{\theta}^j) - V(s_{\theta}^0)) - p (E(s_{\theta}^j, q_{\theta}^j) - E^0(s_{\theta}^0))]}_{\geq 0} \Gamma \underbrace{-T + \epsilon}_{\leq 0}$$

The right inequality is given by Assumption 1(v). The left inequality comes from the following inequalities: $\theta V(s_{\theta}^0) - pE^0(s_{\theta}^0) \leq \theta V(s_{\theta}^0) - pE(s_{\theta}^0, q_{\theta}^j) \leq \theta V(s_{\theta}^j) - pE(s_{\theta}^j, q_{\theta}^j)$. The former is due to technological assumptions about E and E^0 and the latter is due to s_{θ}^j maximizing U .

According to Lemma 2, equilibrium actions s_{θ}^j and q_{θ}^j 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. Three cases arise:

- If $\lim_{\theta \rightarrow 0} NPV(\theta) \geq 0$ then all homeowners participate. Participation is given by $N^j \equiv \int_0^{+\infty} dF(\theta) = 1$.
- If $\lim_{\theta \rightarrow 0} NPV(\theta) < 0$ and $\lim_{\theta \rightarrow +\infty} NPV(\theta) > 0$ then by Equation 14, there exists a unique cutoff type θ_0 such that $NPV^j(\theta_0) = 0$.
- If $\lim_{\theta \rightarrow 0} NPV(\theta) < 0$ and $\lim_{\theta \rightarrow +\infty} NPV(\theta) \leq 0$ then participation is nil. In this case, the gross utility gains accruing to the homeowner never offset the increase in the payment to the contractor.

In what follows, we are interested in the most relevant case where the cutoff type is uniquely defined. Assuming that $F(\cdot)$ is the cumulative distribution function of θ , participation to investment N^j is given by:

$$(16) \quad N^j \equiv 1 - F(\theta_0^j)$$

Finally, under zero-profit condition, aggregate welfare is the sum of utility before investment for non-participating homeowners ($0 \leq \theta < \theta_0$), plus the utility after investment for participants ($\theta \geq \theta_0$):

$$(17) \quad W^j \equiv \int_0^{\theta_0^j} U^0(\theta, s_\theta^0) dF(\theta) + \int_{\theta_0^j}^{+\infty} U(\theta, s_\theta^j, q_\theta^j) dF(\theta)$$

We can now provide a definition of the energy efficiency gap:

Proposition 2. *Assuming that condition (11) is satisfied for all homeowners with $\theta > 0$:*

- (i) *the market is thinner under asymmetric information: $N^{AI} \leq N^{PI}$*
- (ii) *aggregate welfare is lower under asymmetric information : $W^{AI} \leq W^{PI}$*

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

A2.3. Comparative statics

We discuss below further comparative statics with respect to a composite indicator of all market and behavioral parameters: $\zeta \equiv p\Gamma(r, l)$. Any value of p , r or l that does not reflect perfect competition, perfect rationality or perfect information translates into a biased ζ . Specifically, a higher ζ is equivalent to a higher energy price p or a higher Γ , that is, a lower discount rate r or a longer lifetime l . Comparative statics of ζ thus provides insight into the interaction between moral hazard and other market failures or behavioral anomalies.

Applying the Implicit Function Theorem to Equations 6 and 7, we see that an increase in ζ shifts reaction functions $s_\theta^*(\cdot)$ downward and $q^*(\cdot)$ upward:

$$(18) \quad \forall t \quad \frac{ds_{\theta}^*}{d\zeta} = \frac{E_s}{\theta V''/\zeta - E_{ss}} < 0$$

$$(19) \quad \frac{dq^*}{d\zeta} = \frac{-E_q}{C''/\zeta + E_{qq}} > 0$$

This is illustrated in Figure A1 with energy-use externalities.

By the same reasoning as in Lemma 2, a higher ζ entails a higher energy service under asymmetric information. But optimal actions cannot be compared unambiguously under perfect and asymmetric information.

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

$$(20) \quad \frac{dNPV^*}{d\zeta} = - [E(s_{\theta}^*, q_{\theta}^*) - E^0(s_{\theta}^0)]$$

As long as energy efficiency investments decrease energy use for all homeowners, the net present value is increasing in ζ . By the same type of reasoning as in Lemma 2, this leads to a higher participation and a higher average welfare. This conclusion is reversed if all homeowners are subject to an investment-induced backfire rebound effect, i.e., $\forall \theta E(s_{\theta}^*, q_{\theta}^*) > E^0(s_{\theta}^0)$. In this case, a higher ζ decreases participation and average welfare.

A3. Policy solutions

A3.1. Energy-savings contracts and double moral hazard

In our simple framework with no risk-aversion, insurance can be modeled as a contract in which the contractor bears a share k of energy expenditures:

$$(21) \quad U(\theta, s, q) \equiv [\theta V(s) - (1 - k)pE(s, q)]\Gamma - T - I + \epsilon$$

$$(22) \quad \Pi(q) \equiv I + T - C(q) - kpE(s, q)\Gamma$$

According to Assumption 3, the payment to the contractor is $T + I$, where $I = k pE(s, q)\Gamma$ is the actuarially-fair insurance premium.

A new, opposite principal-agent relationship superimposes to the previous one: Since the contractor now provides insurance, he is a principal and the homeowner 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 insurance coverage k and his or her belief about the other party's action. First-order conditions for maximization are:

$$(23) \quad \forall t \quad \theta V' \begin{cases} = (1 - k)pE_s & \text{if } s_\theta^i(k) > s_{min} \\ < (1 - k)pE_s & \text{otherwise} \end{cases}$$

$$(24) \quad C' \begin{cases} = -kpE_q\Gamma & \text{if } q_\theta^i(k) > q_{min} \\ > -kpE_q\Gamma & \text{otherwise} \end{cases}$$

The optimal homeowner's response is bounded above by a satiation value s_{max} .⁵ By the Implicit Function Theorem, the insurance reaction functions $s_\theta^{**}(q, k)$ and $q^{**}(s, k)$ are both increasing in k :

$$(25) \quad \forall t \quad \frac{ds_\theta^{**}}{dk} = \frac{-pE_s}{\theta V'' - (1 - k)pE_{ss}} > 0$$

$$(26) \quad \frac{dq^{**}}{dk} = \frac{-pE_q}{C''/\Gamma + kpE_{qq}} > 0$$

The implementation of such a contract partly solves the moral hazard, as it induces the contractor to offer some quality (Equation 24). 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 use more energy. The energy service in Equation 23 is used to the socially optimal level defined by Equation 6 when the homeowner is not insured ($k = 0$), whereas the quality in Equation 10 is offered to the socially optimal level defined by Equation 7

⁵Without satiation, full insurance ($k = 1$) would bring the marginal value of energy service in Equation 23 to zero, hence inducing infinite energy service. Satiation could be introduced as the argument of the maximum of a parabolic utility function. Alternatively, in our model, it 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.

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 $k \in (0, 1)$. For any insurance k , the agreement $(s_\theta^i(k), q_\theta^i(k))$ 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 perfect-information equilibrium is ambiguous.

Proposition 3. *Optimal energy-savings insurance stipulates incomplete coverage: $0 < \hat{k}_\theta < 1$.*

Proof. See text above. □

Various insurance contracts are illustrated in Figure A4.

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

$$(27) \quad \forall \theta \quad \hat{k}_\theta = \arg \max_{k \in [0,1]} [U(\theta, s_\theta^i(k), q_\theta^i(k)) + \Pi(q_\theta^i(k))]$$

In the second stage, the first-order condition for finding the optimal insurance contract from Equation 27 is:

$$(28) \quad p\Gamma \left(\frac{ds_\theta^{**}}{dk} [\theta V' - pE_s] - \frac{dq^{**}}{dk} \left[\frac{C'}{\Gamma} + pE_q \right] \right) = 0$$

Plugging in Equations 23 and 24 and further rearranging gives the equation that solves the optimal coverage \hat{k} :

$$(29) \quad \forall t \quad p\Gamma \left(kE_s \frac{ds_\theta^{**}}{dk} + (1-k)E_q \frac{dq^{**}}{dk} \right) = 0$$

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.

A3.2. Minimum quality standard

The optimal minimum standard will be set at a value \bar{q} that maximizes collective surplus, net of monitoring cost $M(\bar{q})$, subject to the participation constraint:

$$(30) \quad \begin{aligned} & \text{Maximize}_{\bar{q}} \quad \left[\int_0^{\theta_0} U^0(\theta, s_\theta^0) dF(\theta) + \int_{\theta_0}^{+\infty} [U(\theta, s_\theta^*(\bar{q}), \bar{q}) - M(\bar{q})] dF(\theta) \right] \\ & \text{subject to} \quad NPV(\theta_0, s_{\theta_0}^*(\bar{q}), \bar{q}) - M(\bar{q}) \geq 0 \end{aligned}$$

Assuming that the cutoff type exists and is unique, the constraint in Equation 30 is binding. The optimization 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:

$$(31) \quad \frac{d\theta_0}{d\bar{q}} \left(U^0(\theta_0(\bar{q}), s_{\theta_0}^0) - U(\theta_0(\bar{q}), s_{\theta_0}^*(\bar{q}), \bar{q}) + M(\bar{q}) \right) + \int_{\theta_0(\bar{q})}^{+\infty} \left[\frac{\partial U(\theta, s_\theta^*(\bar{q}), \bar{q})}{\partial \bar{q}} - M' \right] dF(\theta) = 0$$

Recognizing that $U^0(\theta_0(\bar{q}), s_{\theta_0}^0) - U(\theta_0(\bar{q}), s_{\theta_0}^*(\bar{q}), \bar{q}) = -NPV(\theta_0(\bar{q}), s_{\theta_0}^*(\bar{q}), \bar{q})$ and using the binding constraint leads to:

$$(32) \quad \int_{\theta_0}^{+\infty} \left[\frac{\partial U(\theta, s_\theta^*(\bar{q}), \bar{q})}{\partial \bar{q}} - M' \right] dF(\theta) = 0$$

In fact, since marginal participants are indifferent between investing and not investing, marginal changes in participation can be ignored.

In words, the optimal 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.

Note that if participation to investment is nil without the standard, no standard will be welfare-improving. In contrast, if participation is complete without the standard, the constraint will not be binding and the optimal standard will be defined by the following first-order condition: $\int_0^{+\infty} [\partial U(\theta, s_\theta^*(\bar{q}), \bar{q}) / \partial \bar{q} - M'] dF(\theta) = 0$.

A3.3. Interaction between moral hazard and energy-use externalities

Public intervention to address moral hazard problems may not be systematically justified if they interact with energy market failures. Assume that every unit of energy used generates a linear external cost p_x , discounted over the relevant time period with a discount factor Γ_x . For instance, p_x is positive for pollution or energy-security externalities, and negative for average-cost energy pricing. Homeowners' utilities before and after investment is now:

$$(33) \quad \begin{cases} U_x^0(\theta, s) \equiv U^0(\theta, s) - p_x E^0(s) \Gamma_x \\ U_x(\theta, s, q) \equiv U(\theta, s, q) - p_x E(s, q) \Gamma_x \end{cases}$$

These new utility functions allow one to define new net present value NPV_x and aggregate welfare W_x functions as in Equations 5 and 17, respectively. Equilibria are denoted with superscript PE if externalities are internalized through a Pigouvian price p_x and UE if they remain unpriced.

Proposition 4. *In an economy subject to both energy market failures and energy efficiency moral hazard:*

(i) *When energy market failures are corrected, it is desirable to also undo moral hazard problems: $W_x^{PI+PE} \geq W_x^{AI+PE}$*

(ii) *If no homeowner is prone to an investment-induced backfire rebound effect, then it is desirable to correct energy market failures. This holds whether or not moral hazard problems are addressed: $\forall \theta E(s_\theta^{PI}, q_\theta^{PI}) \leq E^0(s_\theta^0) \Rightarrow W_x^{PI+PE} \geq W_x^{PI+UE}$ and $E(s_\theta^{AI}, q_\theta^{AI}) \leq E^0(s_\theta^0) \Rightarrow W_x^{AI+PE} \geq W_x^{AI+UE}$*

(iii) *If homeowners are prone to neither an investment nor a policy-induced backfire rebound effect, then it is desirable to undo moral hazard problems. This holds even if energy market failures are not corrected: $\forall \theta E(s_\theta^{PI}, q_\theta^{PI}) \leq E(s_\theta^{AI}, q_\theta^{AI}) \leq E^0(s_\theta^0) \Rightarrow W_x^{PI+UE} \geq W_x^{AI+UE}$*

Proof. We illustrate with pollution externalities ($p_x > 0$).

(i) For all θ , since $(s_\theta^{PI+PE}, q_\theta^{PI+PE})$ maximizes U_x in the social setting, $U_x(\theta, s_\theta^{PI+PE}, q_\theta^{PI+PE}) \geq U_x(\theta, s, q)$ for all (s, q) , and for $(s_\theta^{AI+PE}, q_\theta^{AI+PE})$ in particular. Likewise, we have $U_x^0(\theta, s_\theta^{0,PE}) \geq U_x^0(\theta, s_\theta^0)$. By Proposition 3, it follows that $W_x^{PI+PE} \geq W_x^{AI+PE}$.

(ii) Again, for all θ , since $(s_\theta^{PI+PE}, q_\theta^{PI+PE})$ maximizes U_x in the social setting, $U_x(\theta, s_\theta^{PI+PE}, q_\theta^{PI+PE}) \geq U_x(\theta, s_\theta^{PI+UE}, q_\theta^{PI+UE})$. In addition, we have $NPV_x(\theta) = NPV(\theta) - p_x \Gamma_x [E(s, q) - E^0(s)]$. Assume θ_0^{PE} is the cutoff type in an equilibrium where both energy-use externalities and moral hazard are addressed, while θ_0^{UE} is the cutoff type in an equilibrium where only moral hazard problems are addressed. We have $NPV_x(\theta_0^{PE}) = 0 = NPV(\theta_0^{UE})$. In the absence of

an investment-induced backfire rebound effect, we thus have $NPV_x(\theta_0^{PE}) = 0 \leq NPV_x(\theta_0^{UE})$. Since NPV is increasing in θ , $\theta_0^{PE} \leq \theta_0^{UE}$, that is, participation is higher if externalities are internalized. The difference in aggregate welfare between the two equilibria is $\Delta W = \int_0^{\theta_0^{PE}} \Delta U_x^0 dF(\theta) + \int_{\theta_0^{PE}}^{\theta_0^{UE}} [U_x(\theta, s_\theta^{PE}, q_\theta^{PE}) - U_x^0(\theta, s^0)] dF(\theta) + \int_{\theta_0^{UE}}^{+\infty} \Delta U_x dF(\theta)$. The first and third integrands of the right-hand side are positive (see proof (i) just above). The second integrand is also positive, since $\forall \theta \geq \theta_0^{PE} U_x(\theta, s_\theta^{PE}, q_\theta^{PE}) \geq U_x^0(\theta, s^{0,PE}) \geq U_x^0(\theta, s^0)$. Therefore, aggregate welfare is larger when externalities are internalized: $W_x^{PI+PE} \geq W_x^{PI+UE}$. The exact same reasoning leads to $W_x^{AI+PE} \geq W_x^{AI+UE}$. This is because since $(s_\theta^{AI+PE}, q_\theta^{AI+PE})$ maximizes U_x under asymmetric information, $U_x(\theta, s_\theta^{AI+PE}, q_\theta^{AI+PE})$ is greater than $U_x(\theta, s, q)$ for any other actions s and q determined under asymmetric information, e.g., $(s_\theta^{AI+UE}, q_\theta^{AI+UE})$.

(iii) Externalities are unpriced here but superscript UE is dropped for the sake of parsimony. Assume θ_0^{PI} (resp. θ_c^{AI}) is the cutoff type in the social (resp. private) optimum. From proposition (4i), we have $\theta_0^{PI} \leq \theta_0^{AI}$. Therefore, the aggregate welfare difference between the two situations is $\Delta W_x = \int_{\theta_0^{PI}}^{\theta_0^{AI}} NPV_x(\theta, s_\theta^{PI}, q_\theta^{PI}) dF(\theta) + \int_{\theta_0^{AI}}^{+\infty} [U_x(\theta, s_\theta^{PI}, q_\theta^{PI}) - U_x(\theta, s_\theta^{AI}, q_\theta^{AI})] dF(\theta)$. In the absence of an investment-induced backfire rebound effect, the first term of the right-hand side is positive (see proof (ii) just above). In the absence of a policy-induced backfire rebound effect, the second term of the right-hand side is also positive. To see this, note that $\forall \theta E(s_\theta^{PI}, q_\theta^{PI}) \leq E(s_\theta^{AI}, q_\theta^{AI}) \Rightarrow -p_x \Gamma_x E(s_\theta^{PI}, q_\theta^{PI}) \geq -p_x \Gamma_x E(s_\theta^{AI}, q_\theta^{AI})$. This, added to $U(\theta, s_\theta^{PI}, q_\theta^{PI}) \geq U(\theta, s_\theta^{AI}, q_\theta^{AI})$ (which is given by definition of the maximum) leads to $U_x(\theta, s_\theta^{PI}, q_\theta^{PI}) \geq U_x(\theta, s_\theta^{AI}, q_\theta^{AI})$. To conclude, the aggregate welfare difference is positive: $W_x^{PI} \geq W_x^{AI}$.

□

As long as energy efficiency does not backfire, correcting energy market failures is desirable, regardless of whether or not the contracting parties overcome the moral hazard. Indeed, social welfare cannot be maximized if the parties do not account for the broader distortions associated with their actions. However, the reciprocal needs not be true: If energy market failures are not (or cannot be) corrected, then it might be desirable to maintain, rather than undo, the moral hazard. This can actually occur if energy efficiency backfires. As a result, energy market failures would be larger.

A4. Sufficient statistics

A4.1. Deadweight loss from supply-side moral hazard

We seek to approximate the deadweight loss associated with the quality shortfall caused by the moral hazard: $\Delta_q W \equiv W^{PI} - W^{AI}$. A first step is to examine the marginal welfare change induced by a marginal change in quality. Since marginal participants are indifferent between investing and not investing, we can ignore changes in participation (see optimal standard). Similar envelope conditions allow us to also neglect the benefits from increased heating comfort. Rewriting Equation 32 with $M(\bar{q}) = 0$, the marginal benefits from a higher quality to participating homeowner are thus:

$$(34) \quad \frac{dW}{d\bar{q}} = p\Gamma E_q - C'$$

Integrating between q^{AI} and q^{PI} (with $q^{AI} \leq q^{PI}$ according to Proposition 2i) gives the following approximation for $\Delta_q W$:

$$(35) \quad \overline{\Delta_q W} = -p\Delta_q E(s^{AI}, q)\Gamma - \Delta_q C(q)$$

The error associated with integrating infinitesimal changes is positive and equal to the private benefits from increased heating comfort and the social benefits from increased participation. Indeed, for a participating homeowner, the exact deadweight loss $\Delta_q W$ is:

$$(36) \quad W^{PI} - W^{AI} = \sum_{t=1}^l \left[V(s^{PI}) - V(s^{AI}) - p \left(E(s^{PI}, q^{PI}) - E(s^{AI}, q^{AI}) \right) \right] \delta^t - \left[C(q^{PI}) - C(q^{AI}) \right]$$

We recognize that:

$$(37) \quad \Delta_q W = \overline{\Delta_q W} + \underbrace{\left[V(s^{PI}) - V(s^{AI}) - p \left(E(s^{PI}, q^{PI}) - E(s^{AI}, q^{PI}) \right) \right]}_{\geq 0} \Gamma$$

The term in brackets is positive because s^{PI} maximizes the function $V(\cdot) - E(\cdot, q^{PI})$. Therefore, $\Delta_q W \leq \overline{\Delta_q W}$. In words, $\overline{\Delta_q W}$ provides a lower bound of the exact average deadweight loss.

The formula is quite intuitive. It weighs the cost of quality against its benefits in terms of gross energy savings. This corresponds to a net present value calculation that only takes into account technological information. It does not require knowledge of the utility function for energy service $V(\cdot)$ nor its specific effect on energy use E_s . Therefore, the direct rebound effect can be ignored. Still, the formula contains the key parameters of the market and behavioral environment p , l and r .

A4.2. Marginal welfare effect of a quality standard

The marginal welfare effect of a quality standard is simply given by Equation 35, enhanced with $-M'$ on the right-hand side. Computation of this formula requires knowing the technology E_q , the production cost $C(\cdot)$, and the cost of random post-implementation audits $M(\cdot)$.

A4.3. Marginal welfare effect of energy-savings contracts

The marginal effect of incremental insurance coverage k to the parties willing to engage in the contract is given by Equation 29 (see optimal standard). It can be rewritten as follows:

$$(38) \quad \frac{dW}{dk} = p\Gamma \left(sE_s\eta_k^s - qE_q\eta_{1-k}^q \right)$$

The η terms are the elasticities of each parties' input to the insurance coverage. The elasticity of the rebound effect to insurance completeness, η_k^s , is positive. The elasticity of quality to insurance incompleteness, η_{1-k}^q , is negative.⁶ These elasticities are the key effects an econometrician would need to measure to evaluate the policy.

Again, computation of the formula requires knowing the technology, namely the average marginal effects of inputs on energy use qE_q and sE_s . But interestingly, unlike the standard, the evaluator does not need to have information about cost $C(\cdot)$.

⁶With isoelastic functions, we have $\eta_{1-k}^q = -\eta_k^q$.

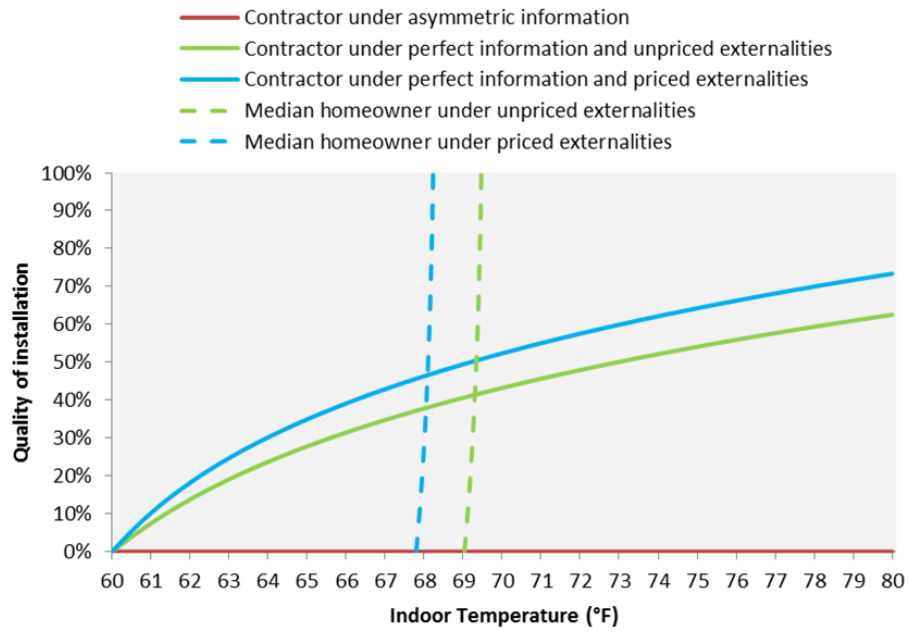


FIGURE A1. Reaction functions, with and without internalization of externalities.

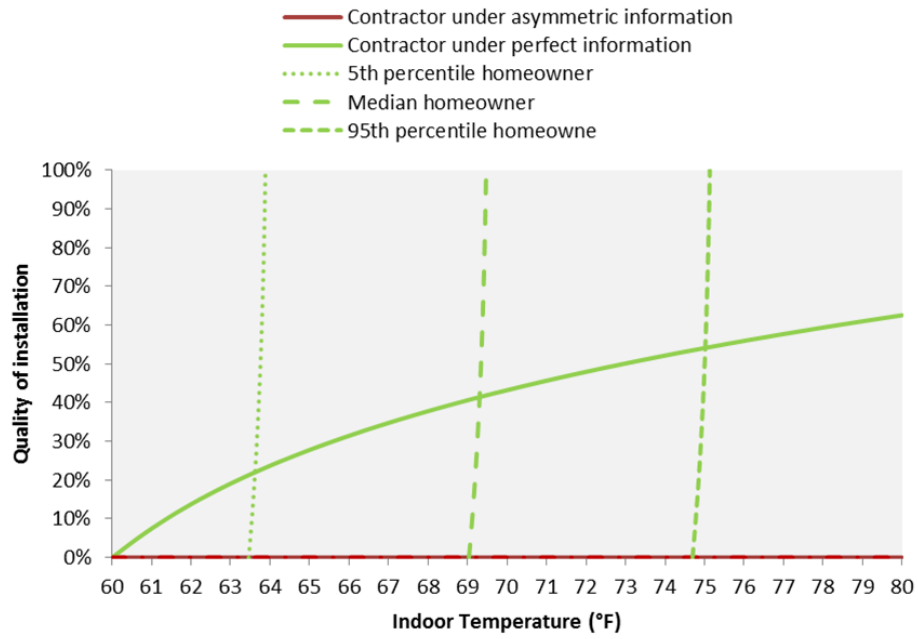


FIGURE A2. Reaction functions with heterogeneous homeowners, without externalities.

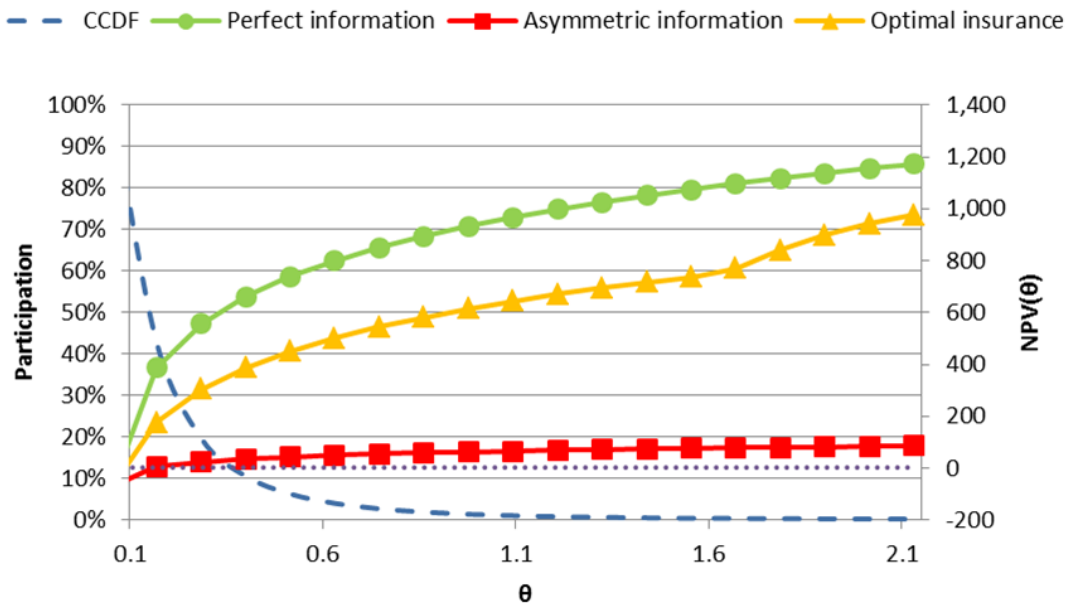


FIGURE A3. **Net present value and participation, without externalities.** The net present value of investment in insulation reads on the right vertical axis. The intersection of each curve with the zero horizontal axis determines the cutoff type θ_0 of the marginal participant in investment. For each cutoff type on the horizontal axis (from the 0.5th to 95.5th 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.

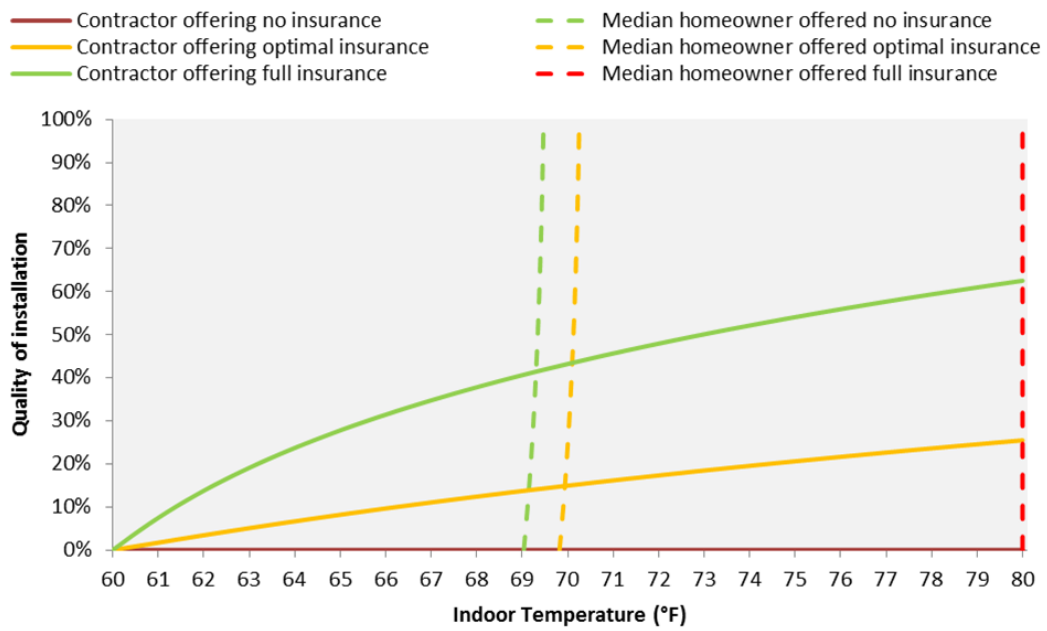


FIGURE A4. **Reaction functions under different energy-savings insurance contracts, without externalities.** The optimal insurance contract displayed here has a coverage of 20%.