NBER WORKING PAPER SERIES

# THE PROBLEM OF THE UNINSURED

Isaac Ehrlich
Yong Yin

Working Paper 18444
http://www.nber.org/papers/w18444

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2012, Revised May 2017

This paper is based on an earlier version presented at "Risk and Choice: A conference in honor of Louis Eeckhoudt", held on July 12-13 in Toulouse, France. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Problem of the Uninsured
Isaac Ehrlich and Yong Yin
NBER Working Paper No. 18444
October 2012, Revised May 2017
JEL No. G22,H42,I13,I28

**ABSTRACT**

The problem of the uninsured cannot be fully understood without considering the role of non-market alternatives to 'market insurance' called 'self-insurance' and 'self-protection' (SISP), including the public 'health care safety-net' system. We tackle the problem by formulating a 'full-insurance' paradigm that accounts for all four interacting insurance measures. We apply two versions of the full-insurance model to estimate, via calibrated simulations, the impacts of SISP on the fraction of uninsured, health spending, and health levels, and to assess how the mandated Affordable Care Act might affect these outcomes in comparison with the CBO projections in 2010. The results indicate that policy analyses which overlook the role of the real price of market insurance relative to the shadow prices of SISP in determining the decision to insure can grossly distort the capacity of mandated reforms like the ACA to insure the uninsured, contain overall health care costs, and improve health and welfare outcomes.

Isaac Ehrlich
415 Fronczak Hall
State University of New York at Buffalo
and Center for Human Capital
Box 601520
Buffalo, NY 14260-1520
and NBER
mgtehrl@buffalo.edu

Yong Yin
415 Fronczak Hall
State University of New York at Buffalo
and Center for Human Capital
Box 601520
Buffalo, NY 14260
yyin@buffalo.edu

***"To insure or not to insure – that is the question"***

## 1. Introduction

The problem of the uninsured – those lacking any health insurance – has generated intense public debate in the United States both before and after the passage into law of the Patient Protection and Affordable Care Act (henceforth ACA), commonly known as Obamacare, in March 2010. The main objective of the ACA has been to induce an estimated 52 million uninsured individuals in a target population of non-elderly, legal resident adults with incomes above 133% of the poverty line to purchase health insurance through the force of regulatory measures and financial incentives.[1] Key among the latter is the individual mandate on consumers to buy insurance, enforceable by income-graduated sanctions[2], and constraints on providers and firms to expand eligibility to, and the coverage by, health insurance policies, while also reducing their costs. The latter is expected to come from creating State-based health insurance exchanges for individuals and small businesses, providing federal subsidies to qualifying groups, and imposing caps on maximum premiums.

The public debate about the desirability of the ACA has been about both the true cost of the mandate and the degree to which it is likely to attain its stated goals. In this paper we try to gain new insights into these issues via two nested versions of a stylized, albeit comprehensive, model of health insurance which mimics the basic features of the new insurance system. The extended version, in particular, accounts for the behavioral implications of an ACA-like mandate and permits an assessment of the mandate's expected costs and benefits via calibrated simulations

A key argument missing in the debate is that being uninsured can be a rational choice even for people who are risk averse, and is not just an outcome of unaffordability. Individuals may choose to eschew market insurance because of its price relative to that of other goods and services, subjective assessments of personal risks, and risk tolerance. Some attention has been devoted to this issue in the literature, especially in the context of insurance against natural hazards (see Kunreuther and Rose 2006), which only a small proportion of the population at risk buys. For example, only 17% of California's homeowners have earthquake insurance.[3] This is also the case for insurance against loss of life. According to a recent survey only 44% of households own an individual life insurance policy; 30% have no individual or employer-provided life insurance; and 11 million households with children younger than 18 have no life insurance.[4] By comparison, less than 20% of the non-elderly adult population lacks health insurance. The possibility that eschewing insurance can be individually optimal is a point of reference in our analysis.

Our more general argument is that the "problem of the uninsured" needs to be assessed as part of a more comprehensive insurance problem which recognizes privately managed alternatives to market insurance as well. These alternatives have been termed "self-insurance" and "self-protection" (see Ehrlich and Becker 1972). Self-insurance refers to actions people take to reduce their magnitude of potential loss from specific harmful hazards, conditional on their occurrence. Self-protection refers to actions individuals take to reduce the probability of the loss occurring in the first place. These individually-controlled measures exist in the case of health hazards as well.

Examples of self-insurance measures that reduce potential health losses are: monitoring one's health conditions to achieve early detection of serious illnesses which ameliorates their severity when illness strikes; improving one's medical literacy to complement health recovery efforts, and acquiring medical savings accounts to reduce the burden of high out-of-pocket costs of health-recovering medical care services, which we call *remedial care.*

Examples of self-protection measures that reduce the likelihood that illness and injury strike are: following a routine of diet, physical exercise, and a myriad of safety measures at work; exercising prudent life-style choices off work; and using preventive medical services, such as annual checkups. The common denominator of these measures is *preventive care*, designed to thwart or reduce potential risks to health, although some preventive measures may reduce both the probability and severity of illness.[5] Market insurance, like self-insurance, serves mainly to limit the significant remedial care costs and health losses incurred if illness strikes.

An alternative to market insurance that is unique to health insurance, however, is the informal "safety net" system. A US federal law known as the Emergency Medical Treatment and Active Labor Act (EMTALA), requires most hospitals to provide emergency care, without consideration of insurance coverage or ability to pay, when a patient arrives at an emergency room for attention to an acute medical condition. Health professional are similarly obliged by the Hippocratic Oath not to deny treatment to people in medical emergencies, and such services are also offered by charitable organizations. This generates a classic ex-ante moral hazard, or "free-rider's" problem, since individuals can take advantage of the system by avoiding payment. The safety net system thus becomes, in principal, a special case of "self-insurance" at zero cost.

Self-insurance is thus intrinsically a substitute for market insurance, although self-protection can in principle be a substitute or a complement, depending largely on whether insurance companies monitor individual efforts at self-protection and reward such behavior with lower premiums – a rather unlikely prospect in the case of typical, menu-based health insurance policies where premiums are based on overall community rating. As we show in Section 2, however, both alternatives, when sufficiently productive, increase the likelihood of a "corner solution" in which insurance is eschewed. This also means that mandating the previously uninsured to purchase insurance can lower individual self-insurance and self-protection efforts. These possibilities, highlighted in Sections 5 and 6, have been largely missing from the debate about the rationale for mandating health coverage, as well as from the micro-simulation models offered by the CBO and Rand's COMPARE, which project the take-up rate of ACA and assess cost implications.

How relevant is this omission empirically and what does it imply about the efficiency of the mandated ACA program? We attempt to answer this question by formulating a "full insurance" paradigm that recognizes the full interaction between market insurance and its alternatives of self-insurance and self-protection (SISP), including the safety-net system, and using it to address the problem of the uninsured via two stylized models: a baseline model in which losses from ill-health are purely monetary and utility is just a function of income, or consumption (Section 3), and an extended model in which utility is enhanced by both consumption and health, and ill health adversely affects both (Section 4). The models are nested in the sense that the first recognizes only *consumption smoothing* as the "full-insurance" objective, whereas the second recognizes both *consumption and health smoothing* as the relevant objectives. Both corroborate the quantitative importance of SISP and enable us to assess quantitatively key intended and unintended outcomes of the mandated ACA system, relative to the pre-ACA system.

The comparison involves four separate behavioral issues we explore in the following sections:

a. To what extent do the specific SISP alternatives to health insurance (including the safety net) account for the *magnitude* of the problem of the uninsured – the percentage uninsured in the target pop – relative to that of the market insurance price or premium (Sections 3 and 4).

b. How important quantitatively are these non-market alternatives in providing "insurance" services as indicated by the extent to which individuals demand them and by their effectiveness in smoothing out income and health fluctuations due to health shocks (Sections 3 and 4).

c. To what extent can the SISP alternatives offset the ACA mandate's effectiveness in achieving compliance, i.e., inducing the uninsured to take up health insurance (Section 5).

d. To what extent do SISP impact the ACA mandate's efficiency in reducing the overall cost of the health care system and in improving the population's health and welfare levels (Section 6).

The stylized full insurance model and its application to health insurance inevitably involve a number of simplifying assumptions. But the model is sufficiently general to allow for calibrated simulations which successfully simulate key empirical data concerning health insurance. The simulations indicate that overlooking the role of SISP as alternatives to market health insurance may grossly overstate the capacity of the ACA mandate to insure the uninsured, contain the overall costs of the health care system, and improve the system's health and welfare outcomes.

## 2. Theoretical Background

The "full insurance" problem incorporates three alternative insurance and protection measures: market insurance (MI), self-insurance (SI) and self-protection (SP), which in turn address three related objectives: consumption-smoothing across different states of the world, loss reduction, and loss prevention. . Consider the binary case having just two relevant states of the health: a “good” state (1) and a “bad” state (0) with endowed probability $p^e$ and loss $L^e$. If the technologies of SI and SP (or SISP), $L(L^e,c)$ and $p(p^e,r)$, aree decreasing and convex functions of their respective opportunity costs c and r, such that $-L'(L^e, c=0)$ and $-p'(p^e, r=0) \rightarrow \infty$, then the SISP alternatives to market insurance, and hence the optimal full-insurance decision encompassing all three (plus the publicly financed health care safety net system), will not be subject to a corner solution. The optimal market insurance component, however, could be nil for people in a heterogeneous population with distinct characteristics including, e.g., endowed odds of illness, income, and attitudes toward risk. This would be the case if the real market price of insurance, $\pi$, representing the terms of trade between income in the good and bad states of the world, is fixed at a level $\pi^0 = [(1+\lambda)p^0/(1-p^0)]$, dictated by the average odds of loss for the “community rated” insurance pool (denoted by superscript 0), $p^0/(1-p^0)$, and a “loading factor”, $\lambda$, which would not reflect differences in individual endowments or efforts at self-protection. This assumption reflects the structure of a typical health insurance policy, which is based on community rating. The fixed price level $\pi^0$ would deviate from its actuarially fair values for most individuals, since it deviates from their varying actuarially fair values.

In the one-period binary case where all potential losses are financial, and income (hence consumption) is the only source of utility, the condition for individual j eschewing market insurance if the latter is the only feasible insurance alternative is given by:

$$(1) \quad \frac{p^e U'(I_0^e)}{(1-p^e)U'(I_1^e)} < (1+\lambda)\frac{p^0}{1-p^0} \equiv \pi^0 ,$$

where $p^e$ is j’s endowed hazard probability; $I_1^e$ and $I_0^e$ are j’s endowed income levels in the “good” and “bad” states the states of world; and $L^e$ is j’s endowed loss, so that $I_0^e = I_1^e - L^e$. The LHS of equation (1) defines the absolute slope of j’s indifference curve between incomes in the good vs. bad states of the world, $UU(p^e)$, assumed to be convex toward the origin[6]. The condition for a corner solution is that $\pi^0$, the slope of the market insurance budget line $MM(p^0)$, cuts the indifference curve from above at the endowment point, E. This is more likely if one’s endowed probability of suffering a loss is low and the community-based insurance loading term is high (see Fig. 1). How would the choice change if self-insurance and self-protection became feasible?

a. *Self-insurance*. An effective convex technology for loss reduction, or more generally income-transfer between states 1 and 0, assumes a shape like the transformation curve $TT_1$ in Figure 1.[7] Self-insurance would always be adopted if the absolute slope of $TT_1$ at point E, $-1/[L'(L^e,c)+1]$, is lower than that of the indifference curve passing through E (not shown in the graph). If self-insurance were the only means of “insurance”, its optimal value, c*, would then be attained at the point of tangency between the indifference curve $UU(p^e)$ and $TT_1$, $S_1$. If the slope ($\pi^0$) of the market insurance budget line passing through $S_1$, $MM(p^0)$ were steeper than that of $TT_1$, however, it would also be steeper at point E because of the convexity of $TT_1$. Self-insurance would then completely "crowd out" market insurance. The condition for j eschewing MI is that

$$(2)\quad -\frac{1}{L'(L^e, c^*)+1} = \frac{p^e U'(I_0^*)}{(1-p^e)U'(I_1^*)} < \pi^0,$$

i.e., the exogenously fixed market insurance terms of trade $\pi^0$ exceed those of self-insurance at the latter's optimal level (see point $S_1$ in Fig. 1).

b. *Self-protection*. Assuming that individual self-protection offers a similarly effective and convex technology for loss prevention, SP too would always be adopted if its initial marginal product is sufficiently high, $-p'(r \rightarrow 0) \rightarrow \infty$, since in this case the shadow price of an increase in self-protection would fall short of the income equivalent of its marginal value in utility [8], and its effect would be manifested as a reduction in the absolute slope of the individual j's indifference curve going through point $S_1$. If the market insurance price remains constant at $\pi^0$, as is the case when health insurance premiums are based on community rating, the market insurance budget line would remain $MM(p^0)$. The slope of the indifference curve at $S_1$ would now become flatter, reflecting a lower probability (p*) and odds of loss generated by self-protection. Equilibrium would shift from $S_1$ to $S_2$ – the new tangency position between $TT_1$ and the highest attainable indifference curve UU(p*), generating a reduction in SI. Market insurance remains nil since the slope of the transformation curve becomes even lower relative to $\pi^0$, i.e.,

$$(3)\quad -\frac{1}{L'(L^e, c^{**})+1} = \frac{p(p^e, r^*)U'(I_0^*)}{[1-p(p^e, r^*)]U'(I_1^*)} < \pi^0,$$

where c** and r* denote the optimal self-insurance and self-protection (SISP) opportunity costs at point $S_2$. This analysis can be summarized by the following propositions:

**Proposition 1.** *Some form of SISP must be adopted – "full insurance" cannot be nil – if both the indifference curves between income in the good vs. bad states of the worlds as well as the technologies of self-insurance and self-protection are strictly convex such that both* $-L'(L^e, c=0)$ *and* $-p'(p^e, r=0) \rightarrow \infty$.

Proof: These conditions guarantee that the productivity of some positive spending on self-insurance would exceed the marginal rate of substitution in utility at any positive value of c and r respectively so that spending at least some positive amounts on SI and SP would always be optimal, even if market insurance were eschewed. "Full insurance", encompassing all three forms of insurance could then never be nil.

**Proposition 2.** *Self-insurance and self-protection are separately and jointly substitutes for market insurance; an improvement in the technologies producing each or both raises the likelihood that market insurance is eschewed when the net price of insurance is fixed by a* *uniform community rating which does not account for endowed individual health risks and the moderating effects of SISP on these risks.*

Proof: Graphically, improvements in the SISP technologies make both the transformation curve and the indifference system flatter, relative to the market insurance budget line. Continuing improvements in these technologies can ultimately make the absolute slopes of both curves lower than that of the market insurance budget line, $\pi^0$, at the endowment position, causing the consumer to avoid purchasing market insurance and preferring an optimal combination of SI and SP, as illustrated in Figure 1.

**Proposition 3.** *For a given optimal amount of market insurance, MI, optimal self-insurance (c*) and self-protection (r*) are substitutes. If the former rises, e.g., the latter falls.*

Proof: This is easily seen if market insurance is nil. In this case, an improvement in the self-protection technology necessarily increases the optimal amount of self-protection and lowers the optimal amount of self-insurance, since the equilibrium position associated with the joint optimal solution for self-insurance and self-protection (SISP) would then move leftward from $S_1$ to $S_2$ on the transformation curve $TT_1$ in Fig. 1.

**Proposition 4.** *For risk averse consumers, the last dollar optimally spent on self-protection (r*), relative to self-insurance (c*), has a larger proportional impact on the probability of loss (p*) compared to the severity of loss (L*), i.e.,* $-\, dlnP/dr^* > -dlnL/dc^*$. *The same holds for the impact of the optimal marginal spending on self-protection relative to self-insurance in reducing one's expected loss, p*L*, and hence one's gross expected income* $I_1^e - p^*L^*$.

Proof (Chang and Ehrlich, 1985). Self-insurance necessarily causes a greater reduction of the variance of income relative to self-protection at a level of expenditure where both cause an equal reduction in expected income. Since for the risk averse SI would then generate a bigger expected utility gain, optimal self-protection would need to yield a bigger absolute reduction in expected income (via a greater percentage reduction in p) relative to self-insurance (via a smaller percentage reduction in L).[9]

Proposition 4 offers a corollary: *preventive care plays a quantitatively bigger role than remedial care in controlling expected losses from ill health.* The corollary holds on the assumption that preventive care is oriented toward avoiding illness (self-protection or loss prevention), while remedial care focuses on health restoration or recovery.

These propositions indicate that self-insurance and self-protection (SISP) may have a non-trivial impact on health insurance choices. A remaining issue, however, is *how important* is this impact quantitatively. We address this issue in the following sections.

## 3. Baseline Model

### A. Simplifying assumptions:

We consider a heterogeneous population stratified by endowed probabilities of sickness, $p^e$, which are uniformly distributed on the open interval (0 1). For simplicity of exposition, all other parameters characterizing potential differences across the heterogeneous risk groups (such as differences in income endowments; efficiency parameters controlling the production of SISP; or insurance premiums set by community ratings) are abstracted from in order to focus on the critical role of heterogeneity in endowed morbidity risks. The loss from getting sick, $L^e$, is purely monetary; we ignore any consumption needs associated with health insurance and take insurance to be an indemnity type. This enables us to abstract from any ex-post moral hazard, or excess consumption of insured medical care, and to focus on the role of SISP in the traditional insurance model. However, both assumptions are relaxed in the extended model we develop in Section 4.

a. *Specifying the menu-type health insurance policy*: The policy offers a fixed indemnity, menu-based insurance coverage with no choice of "partial coverage" by way of varying coinsurance rates or deductibles. The policy sets a single premium level, R, based on community rating.[10] Since the policy is thus inherently *actuarially unfair*, the coverage (payout) rate is restricted not to exhaust the endowed loss, so $L^a \leq L^e$, where $L^a$ sets a maximum level of coverage, accounting also for personal losses such as sick time, which are typically not covered by insurance. Under these conditions, the insurance policy becomes a "take it or leave it" proposition.

b. *Specifying the self-insurance production function*: Self-insurance can be described as lowering the potential illness loss $L^e$ by a proportion A(c), which is a function of SI spending, c, thus allowing $L^e$ to fall to a lower level, $A(c)L^e$. That is, $L(L^e, c) = A(c)L^e$. The production function governing A(c) is specified as the convex function

$$(4)\quad A(c) = A_h + (1 - A_h)\exp(-\eta_1 c),$$

with $A(0) = 1$ and $-A'(c{=}0) \to \infty$, so some SI is always optimal, but $A(\infty) = A_h$, to set a limit on the effectiveness of SI as reflected by the production possibilities frontier $TT_1$ in Figure 1 (see also the discussion of the self-protection production function below).

c. *Specifying the safety net (SN) health care services*: The safety-net care is assumed to be available at zero cost. But it is also assumed to be provided as an inferior "indemnity" – a minimum quality of care limiting the maximum loss coverage to $L^0$, which is significantly below the maximum coverage provided by the insurance policy, i.e.,

$$(5)\quad L^0 << L^a.$$

d. *Specifying the self-protection production function*: The probability of falling ill, like its associated loss, can be lowered by a proportion B(r) of its endowed value, $p^e$, i.e.,

$p(p^e,r) = B(r)p^e$, with B(r) specified as a convex production function of self-protection spending, r, as follows:

$$(6)\quad B(r) = B_h + (1-B_h)\exp(-\eta_2 r),$$

with $0 < B_h <,1$ $B(0) = 1$ and $-B'(r=0) \rightarrow \infty$, setting a minimum level for r*, and $B(\infty) = B_h$ setting a limit on the effectiveness of SP. We choose the same functional specification for the production technologies (other than their idiosyncratic parameters) in equations (4) and (6) for two reasons: a. they are bound by the same limiting constraints, since both self-insurance and self-protection aim to *ameliorate* adverse health outcomes compared to their levels in good states of health; and b. asymmetric specifications need to be defended as special cases. The symmetric specifications thus serve as a reasonable baseline, or "neutral', specification. [11]

e. *Specifying the utility function*: Utility is assumed to be a strictly concave function of income (or consumption) and exhibit constant relative risk aversion (CRRA), commonly used in the literature as follows:

$$(7)\quad U(I) = \frac{I^{1-\sigma}-1}{1-\sigma},$$

with the risk tolerance coefficient calibrated at σ =2, as is conventionally assumed in the literature.

**B. The maximization problem**:

The insurance decision "to buy or not to buy" involves a straightforward decision criterion for j: whether the expected utility associated with buying insurance (IN) exceeds or falls short of the corresponding expected utility associated with staying uninsured (UN)

If the IN option is chosen, the wealth prospect involves the following income distribution:

$I_1^{IN} = I_1^e - R - c - r$, with probability of $1 - B(r)p^e$

$I_0^{IN} = I_0^e - R - c - r - A(c)L^e + L^a$, with probability of $B(r)p^e$

If the UN option is chosen, the wealth prospect involves the alternative income distribution:

$I_1^{IN} = I_1^e - c - r$, with probability of $1 - B(r)p^e$

$I_0^{IN} = I_0^e - c - r - A(c)L^e + L^0$, with probability of $B(r)p^e$.

The expected utility function for both the insured and the uninsured is given by the general form:

$$(8)\quad EU^N(c,r) = B(r)p^e U(I_0^N) + [1 - B(r)p^e]U(I_1^N),$$

where the superscript N in (8) stands for both the insured (IN) and the uninsured (UN), respectively. Members of both groups would then choose optimal levels of SI and SP (c and r) to maximize their expected utility, which satisfy the first-order conditions:

$$(9) \qquad A'(c^*) = -\frac{1}{L^e}[1 + \frac{1 - B(r^*)p^e}{B(r^*)p^e}\frac{U'(I_1^N)}{U'(I_0^N)}],$$

$$(10) \qquad B'(r^*) = -\frac{B(r^*)p^e U'(I_0^N) + [1 - B(r^*)p^e]U'(I_1^N)}{p^e[U(I_0^N) - U(I_1^N)]},$$

For N = (IN, UN). Let, now the maximized value of equation (8) for the insured and uninsured be denoted $EU^{*IN}(c^*, r^*)$ and $EU^{*UN}(c^*, r^*)$, respectively, where c* and r* are the solutions for optimal SI and SP expenditures by the insured and the uninsured, respectively. Then, individual j would purchase health insurance if, and only if

$$(11) \qquad EU^{*IN} \geq EU^{*UN},$$

but would stay uninsured otherwise.

### C. Calibration

We calibrate both the simplified baseline model and the expanded model in Section 4 on data for non-institutional US legal citizens who are nonelderly adults, and live in households with income higher than 133% of the federal poverty line (see fn. 1). The main source of data for this target population is the 2009 Medical Expenditure Panel Survey (MEPS).

As stated earlier, we set $\sigma = 2$ in the CRRA function we selected for the model (see equation 7), as commonly assumed in the literature. Most of the other parameters are taken from information provided in the MEPS 2009 data. We set income at $36,000 to match the average income of the target population in MEPS. The premium R is set to be $960, based on the average employee's contribution in MEPS. As for the free parameters of the production functions for SISP, as specified in equations (4) and (6), we select $\{A_h = 0.8, \eta_1 = 0.05\}$ and $\{B_h = 0.7, \eta_2 = 0.05\}$, respectively. According to MEPS the average (expected) level of medical expenditure is $3,300. We set the maximum indemnity coverage to be $L^a = 0.8L^e$ (to limit the loss-restoring capacity of SI). The endowed loss, $L^e$, is calibrated via our simulations to be $8,250. We therefore set $L^a$ to be $6,600.

The remaining free parameter in our simulation – the coverage level $L^0$ provided by the safety net system – is calibrated to match the fraction of the uninsured population - assessed in the 2009 MEPS to be 20% of the target population. This yields the value of $L^0 = .1273L^e = \$1050$.

### D. Solving the model

In solving the model, we aim to achieve the following objectives: a. assessing the degree to which each of the four components of the full-insurance choice can account for the decision to eschew insurance; b. estimating numerically the optimal *demand* for SISP as well as the latter's impact on the probability and severity of losses from ill-health. One advantage of this model, unlike the extended model we develop in Section 4, is that it has closed form solutions which are consistent with the propositions of Section 2.

**a.** *Assessing the influence of the four insurance alternatives on eschewing insurance*

Using calibrated simulations, we have been able to match the official estimate of the fraction of the target population that is uninsured at 20% and to derive upper bounds for the degree to which the alternative components of the full-insurance insurance account quantitatively for the non-insurance decision. That is, of the 20% uninsured we estimate that up to 50.3% are motivated by the availability of the three non-market alternatives of insurance: the safety net system, which motivates 3.8% [20-16.2] of the target population or 19% (3.8/20) of the uninsured, and the combined alternatives of self-insurance and self-protection, which motivate 6.25% of the target population [20-13.75], or 31.3% of the uninsured [6.25/20]. At least 49.7% of uninsured are thus assessed to eschew insurance because of the actuarially "unfair" price of market insurance (see Table 1).[12]

b. *Quantifying the relative demand for self-insurance and self-protection as components of full insurance and their impact on the prospective loss from illness.*

We also use our calibrated simulations to quantify the demand for, or optimal spending on SI and SP, c* and r*, in both absolute terms and relative to the premium for health insurance under the alternative scenarios of being insured or uninsured at varying endowed probabilities of illness. Table 2 shows the results.

First, consistent with Proposition 2 in Section 2, SI and SP are shown to be substitutes for market insurance: the quantitative values of c* and r* are consistently larger for the uninsured relative to the insured. Indeed, while outlays on SI by the uninsured exceed those by the insured by an average of 10%, the outlays on self-protection by the uninsured exceed those by the insured by 230% on average.

Second, optimal SI and SP also vary by the magnitude of the endowed risks of illness, but there is again a generally significant difference in this regard between the insured and the uninsured: when the endowed probability of ill health rises from 0.1 to 0.5, SI rises by 83% for the insured group, while the magnitude of SP spending hardly varies over the same range. The pattern is consistent with the inherent role of SI as a substitute for market insurance (which does not vary in magnitude in the baseline model), as well as with the role of self-protection, which can be either a substitute or a complement for market insurance as well. For the uninsured, however, both SI and SP serve as substitutes for the absent market insurance. Indeed, both c* and r* rise by about 70% when the endowed risk of illness rises from 0.1 to 0.5.

Third, despite these somewhat different patterns for SI and SP, however, the overall spending on both rises continuously over the same range of endowed probabilities of illness. Computed as a percentage of the employees' contribution to the premiums charged for health insurance, which sets the effective premium as R=$960, the combined outlays $(c^*+r^*)/R$ rise from 6.07% to 9.39% for the insured, and from 9.19% to 15.32% for the uninsured, when $p^e$ rises from 0.1 to 0.5.

Our calibrated simulations of the baseline model also enable us to assess numerically the projected impacts of SI and SP on the endowed magnitudes, probabilities, and expected values of the prospective losses from illness. Table 3 shows the estimated values of these reductions. As indicated by the ratios of the optimized relative to the endowed magnitudes, all are larger for the uninsured group relative to the insured group, since in the baseline model, market insurance is a substitute for both SI and SP. Furthermore, the impact of optimal self-protection (the values of $p^*/p^e$) are larger than those of optimal self-insurance, in conformity with Proposition 4 in Section 2. The same holds for the relative magnitudes of the expected losses from illness. The larger SI and SP efforts by the uninsured, as shown in Table 2, are shown in Table 3 to result in smaller expected income losses than the corresponding SI and SP efforts by the insured, especially at higher levels of endowed probability of illness, both within each group and across groups under the same endowed risks of illness.

The calibrated simulations also verify that SI and SP are substitutes, in line with Proposition 3 in Section 2. We calculate that the optimal amounts of c for the insured (at an endowed loss probability of 0.24, e.g.) and the uninsured (at an endowed probability of 0.14) are respectively $55.61 and $48.51. We then calculate that if self-protection is not available, the amounts of c at the same endowed probabilities would be $59.79 and $54.67, respectively. This implies that when self-protection is made available to an insured person, optimal spending on self-insurance would decrease by $6.61, or 11.3%. The drop would be $4.18, or 7%, for the insured.

c. *Assessing the impact of the insurance price and other parameters on the insurance decision*

The calibrated simulations of the baseline model also provide insights about the implicit role of the unit price of insurance in explaining the problem of the uninsured. The latter is specified in Section 2 to equal $\pi = (1+\lambda)p/(1-p)$, where $\lambda$ denotes the insurance loading factor which indicate the deviation of the actual insurance price from its actuarially fair value $p/(1-p)$. The price, $\pi$, can also be shown equal to (coverage – premium)/premium (see Ehrlich and Becker 1972). Using MEPS data about average insurance coverage and effective premium in the target population we calibrate $\pi^0 = (6600-960)/960 = 5.875$. This analysis indicates the degree of actual unfairness (gross loading) represented by the uniform premium for people with varying personal endowed probabilities of ill health. At $p^e = 0.1$, the uniform price represents a gross loading of $(1+\lambda)=$ 52.875, while at $p^e = 0.8$, the gross loading is just 1.468. The relative gross loading term imposed on the lowest risk group in the population is thus potentially 36 times higher than that imposed on the highest risk group, by these estimates.

Indeed, we can recover the critical value of the endowed probability of ill-health (the source of individual heterogeneity) which is associated with the separating equilibrium in the model, i.e., the point at which the population would split between those choosing to be insured as opposed to being uninsured. That value is estimated to be $(p_0^e)^*$=B(r*)$p^e$ = .143. That is, those with optimized probabilities of ill-health lower than 14.3% will choose to be uninsured. This result is just illustrative since the simplified baseline model assumes that p is the only source of heterogeneity in the target population. It nevertheless suggests that the relatively healthy have a strong incentive to opt out of the insurance market when the price of insurance is uniform and thus relatively more unfair to them than to the average individual.

The behavioral implication emanating from this analysis is that individuals with, say, a 10% endowed probability of loss due to ill-health would optimally choose to be uninsured. Our calibrated simulations thus suggest that if forced to be insured, these individuals would reduce their self-insurance spending by 9.53% and their self-protection spending by 56.48%, as seen from row 1 of Table 2.

As for the role of the other major determinants of the full insurance decision, see our analysis of comparative statics under the decision to accept or reject insurance (see Appendix A.1).

## 4. Extended Model

### A. Relaxing key limiting assumptions:

The baseline model enables an assessment of the role of the full insurance decision under the conventional insurance framework, where smoothing income fluctuations due to stochastic medical care needs is the exclusive goal. The implicit assumption is that the goods on which income is spent are all perfect substitutes. The extension we develop in this section recognizes health and ordinary consumption to be distinct, but complementary goods, and "health smoothing" to be the basic objective of health insurance. The latter can be achieved through insurable "remedial care" services that help restore health loss in the "bad" state of the world when illness strikes. This extension thus requires the specification of a new production function, linking health restoration to remedial medical care services that are covered by health insurance, but can also be financed via out-of-pocket payments. It also exposes the role that a typical health insurance policy plays in affecting one's chosen level of health care services: by allowing health insurance to reimburse consumers for their actual health spending under a fixed premium, as is typically the case, market insurance can generate ex-post moral hazard or "overconsumption" of medical services, which the baseline model abstracts from by assuming indemnity-type insurance. How would the extended model affect our assessment of SISP's relevance for the problem of the uninsured?

To answer these questions we first modify the CES utility function in equation (7) as follows:

$$\text{(7a)} \qquad U(H,X)=\frac{(H^{\theta}+X^{\theta})^{\frac{1-\sigma}{\theta}}-1}{1-\sigma},$$

where H denotes health, or health benefits, and X denotes ordinary consumption. We specify $\sigma = 2$ as in equation (7) but allow $\theta$ – the degree of complementarity between H and X, constrained to be between $-\infty$ and 1 – to be determined by our calibration analysis. Note that $\theta = 1$ would leave X and H to be perfect substitutes, as (implicitly) in the baseline model, whereas $\theta = -\infty$ would make them perfect complements.

In this specification, whether health and consumption are complements or substitutes in utility would thus depend on the *cross derivative* of utility with respect to H and X, $U_{HX} \equiv \partial^2 U / \partial H \partial X$. Under our assumed $\sigma = 2$, it is easy to see that if $\theta > -1$, then $U_{HX} < 0$ and two would be substitutes in utility. The converse holds if $\theta < -1$. Formally, both possibilities are admissible as a matter of idiosyncratic preferences, but complementarity in utility is more defensible in terms of an efficiency principle since better health, as a special form of human capital, enhances the degree of satisfaction one can derive from most "consumption activities", which require that the consumer is in good physical and mental health to fully enjoy them. In our calibrated simulation analysis, $\theta$ is treated as an open parameter to be determined by the calibrated numerical simulation. The calibrated value we will obtain is $\theta = -7$, confirming our expectation that H and X are complementary in utility.

The technologies governing self-insurance and self-protection in the extended model – offering essentially "preventive care" services by lowering the endowed probability and severity of the health losses if illness strikes through self-efforts – remain the same as those we used in the baseline model (see equations 4 and 6) - based on the same arguments we have used to rationalize their symmetrical specification, since these apply equally in the extended model as well. Self-financed spending on "remedial medical care", however, can be done via market insurance and/or out-of-pocket payments as well.

a. *Specifying the opportunities to remedy health losses via medical care*: If illness strikes, endowed health, $H^e$, is subject to a potential health loss of $L^e$. The *potential* loss can be reduced via preventive self-insurance to a level $A(c)L^e$, as in the baseline model. But the *actual loss* can be remedied, or restored, via medical care services, M, that are available at a *relative* price Pm (the price of ordinary consumption, X, being the numeraire), which can remedy the lost health by an amount $L^a\varphi(M)$, where $L^a < L^e$ is the loss control limit reachable via remedial care.[13] The production function linking the fraction of restored health loss to remedial care is given by

$$\text{(12)} \quad \varphi(M) = 1 - \exp(-\eta_3 M),$$

This production function has the property that $\varphi(0) = 0$ and $\varphi(\infty) = 1$.

b. *Specifying the health insurance policy*: As is typically the case, the health insurance provider reimburses policy holders for their medical care at a fixed coinsurance rate of $0 < \kappa < 1$without a cap on spending. The policy sets a single premium level, R, as in the baseline model. The choice of whether to insure or not to insure thus remains a "take it or leave it" proposition.

c. *Specifying the safety net (SN) health services*: We continue to allow for safety-net health-care services to be available at zero cost, albeit as an inferior "indemnity" – a minimum quality of care limiting the maximum recovered loss to $L^0 << L^a$ for the non-insured. But we also allow the uninsured to purchase additional health care services out of pocket, to further reduce their health losses up to the maximal reduction of $(L^a - L^0)$, using the same technology as the insured.

B. **The maximization problem**:

The insurance decision "to buy or not to buy" again involves a straightforward decision rule for individuals: whether the expected utility associated with buying insurance (IN) exceeds or falls short of the corresponding expected utility associated with staying uninsured (UN). There are 4 control variables to select: remedial care, M, consumption, X, self-insurance, c, and self-protection, r. The overall optimization problem can be characterized heuristically as a two-step procedure: in the "first", one selects utility-maximizing levels of M and X that are conditional on given levels of c and r. In the "second" one chooses the utility-maximizing levels of c and r subject to one's optimally chosen schedules of M and X. In each step, a further distinction needs to be made, conditional on whether the individual winds up choosing to be insured or uninsured. With all conditional choices settled, one can finally also settle the ultimate decision whether to insure or not to insure. In reality, all of these choices are made simultaneously.

Step 1: Solving for optimal M and X given c and r

a. If the Insurance option is chosen, the health level in the state of sickness (0) would be given by $H_0^{IN} = H^e - A(c)L^e + L^a\varphi(M)$. One maximizes the utility function (7a) with respect to M and X subject to the budget constraint,

$$(13)\quad \kappa P_m M + X + c + r + R = I_0^e,$$

and the health production function (12). The optimal values of M and X in state 0 must satisfy:

$$(14)\quad M^* = \frac{\log(L^a) + \log(\eta_3) - \log(P_m) - \log(\kappa) - (1-\theta)\log\left(\dfrac{H^*}{X^*}\right)}{\eta_3}$$

$$(15)\quad X^* = I_0^e - \kappa P_m M^* - c - r - R\ .$$

Based on these values, we label the conditionally maximized utility level in the state of sickness

(16) $U^{IN}(H^*, X^* \mid c, r) \equiv U_0^{IN}$.

If the state of good health (1) occurs, the utility level of the insured can be denoted simply

(17) $U^{IN}(H^e, I_1^e - c - r - R) \equiv U_1^{IN}$.

b. If the no-insurance option is chosen, the health level in the state of sickness (0) would be: $H_0^{UN} = H^e - A(c)L^e + L^0 + (L^a - L^0)\varphi(M)$. One maximizes the utility function (7a) with respect to M and X subject to the budget constraint:

(13a) $P_m M + X + c + r = I_0^e$

and the health production function (12). The conditions for optimal M and X in state 0 are then

$$(18)\ M^* = \frac{\log(L^a - L^0) + \log(\eta_3) - \log(P_m) - (1-\theta)\log\left(\frac{H^*}{X^*}\right)}{\eta_3}$$

(19) $X^* = I_0^e - P_m M^* - c - r$.

In this case the analogs to equations (16) and (17) would be:

(16a) $U^{UN}(H^*, X^* \mid c, r) \equiv U_0^{UN}$, and (17a) $U^{UN}(H^e, I_1^e - c - r) \equiv U_1^{UN}$

Step 2: Optimizing on c and r given M and X

Using the conditionally maximized utilities reflecting the optimal schedules of M and X in equations (14) and (15), we can now specify the expected utility function to be maximized with respect to c and r if one chooses to be either insured or uninsured by the general form:

(8a) $EU^N(c, r) = B(r)p^e U_0^N + [1 - B(r)p^e]U_1^N$,

where N stands for both IN and UN. The expected-utility-maximizing values of SI and SP (c* and r*) must satisfy the first-order conditions:

$$(20)\ A'(c^*) = -\frac{1}{L^e}\left[\frac{\frac{1-B(r^*)p^e}{B(r^*)p^e}U'^N_{1X} - U'^N_{0X}}{U'^N_{0H}} - \frac{L^a\varphi'}{P_m}\right],$$

$$(21)\ B'(r^*) = -\frac{B(r^*)p^e(U'^N_{0H}\varphi'/p_m + U'^N_{0X}) + [1 - B(r^*)p^e]U'^N_{1X}}{p^e[U_0^N - U_1^N]},$$

where the subscript s stands for the state of the world, i.e., s= {0,1}, and $U'^N_{sH} = \partial U^N_s(H,X)/\partial H$, e.g., denotes the partial derivative of $U^N_s$ with respect to H.

We proceed by solving for the unconditionally maximized value of equation (8a) for the option of being insured or uninsured $EU^{*IN}(c^*,r^*)$ and $EU^{*UN}(c^*,r^*)$, respectively. Specifically, individuals would choose to be insured if and only if

(22) $EU^{*IN} \geq EU^{*UN}$.

## C. Calibration

In applying the extended model we adopt for most of the parameters the same values we used in the baseline model. The list includes income in the good state, $I_1^e$, premium, R, and the parameters σ, $A_h$, $B_{h,}$, $\eta_1$, and $\eta_2$ defining the utility and production functions of SISP. In addition, we set the coinsurance rate at 25% (cf. Manning and Marquis, 1996), and Pm at 1.75 - the ratio of the medical CPI to the general CPI in 2009. We then calibrate through the numerical simulation the joint set of free parameters defining the health production and utility functions $\eta_3$, and θ, respectively, the endowed levels of health and health loss, $H^e$ and $L^e$, and the safety net level of recovered health loss, $L^0$ to match the average (theoretically, the expected) level of individual medical expenditures ($3,300) and the percentage of the uninsured non-elderly adults in the population (20%), as reported in the 2009 MEPS.

## D. Solving the model

Following the outline we used to report the numerical results of the baseline model, we focus below mainly on the similarities and differences in the results we obtain from applying the extended, relative to the baseline model.

a**.** *Assessing the influence of the four insurance alternatives on eschewing insurance* (Table 4)

Table 5 reports a very similar breakdown of the population of uninsured. Of the 20% uninsured in the target population we estimate that 45.5% are motivated by the availability of the three alternative forms of insurance: the safety net system, which motivates 3.51% [20-16.49] of the target population or 17.6% (3.51/20) of the uninsured; and the combined alternatives of self-insurance and self-protection, which motivates 5.64% of the target population [20-14.36], or 28.2% of the uninsured [5.64/20]. The remaining 54.5% of uninsured can be explained by highly "unfair" price of market insurance as viewed by the uninsured. The main difference in the results obtained from the extended, relative to the baseline model is that the percentage of the uninsured motivated by the three alternative measures is lower (45.5% relative to 50.3%), which implies that a larger fraction was motivated by the high price of market insurance, essentially because the more realistic insurance contract enables coverage of a chosen level of remedial medical outlays and is not restricted by a fixed indemnity.[14]

b. *Quantifying the relative demand for and impact of self-insurance and self-protection*

The pattern of the results as seen in Table 5 is similar to that summarized in Table 2, but the magnitudes of outlays on SI and SP become considerably higher in absolute terms and as percentages of the premium in the extended model relative to the baseline model. This is seen especially in the case of self-insurance by both the insured and the uninsured, but also in the case of self-protection, especially by the insured. The combined spending on SI and SP relative to market insurance doubles for the insured but also rises for the uninsured in Table 5 relative to 2.

Spending is still consistently larger under the option of being uninsured relative to being insured, but the differences are now narrower: only 2% in the case of SI but much higher in the case of SP where the increases range from 43% to 102%. Spending on both SI and SP is also seen to rise when the probability of incurring a loss rises from 10% to 50%.

Clearly, the main reason for the higher spending on all forms of insurance under the extended model relative to the baseline model is the recognition of health as a distinct but complementary commodity to ordinary consumption in the extended, relative to the baseline model. This increases the motivation to reduce the probability and severity of health losses. Indeed, the higher spending on SI and SP leads to a greater reduction in the magnitudes of the probability ($p^*/p^e$) and severity ($L^*/L^e$) of the loss in Table 6 relative to Table 3, with the impact remaining more pronounced for the uninsured. As is the case in table 3, Table 6 also indicates that the percentage fall in p* is larger than that in L*, confirming Proposition 4 in Section 2.[15]

Table 5 also shows that under the reimbursement-type insurance we allow for in the extended model, medical care spending substantially exceeds that on SI and SP, especially at higher levels of endowed illness probabilities. The demand is lowest at probability levels of 10% and 20% where consumers are optimally uninsured, but becomes much higher at probability levels higher than 20% where consumers are optimally insured. Under both options, medical spending would rise by 361% and 389%, respectively, when the endowed risks of illness rise from 0.1 to 0.5.

c. *Quantifying the role of other key determinants of the full insurance decision*

As for the role of other determinants of the problem of the uninsured, the value of the critical probability of loss which produces the separating equilibrium concerning the choice of being insured rather than uninsured is $p_1^{*e}$ =14.3% - practically identical to its value in the baseline model. In addition, the calibrated uniform price of insurance $\pi^0$ and its varying gross loading terms remain *identical* to those derived in the baseline model. The comparative static effects of the extended model's major control variables also remain virtually the same as in the baseline model, except that here we can also estimate their impact on optimal medical care spending as well (see Table A1 and Table A2 in the Appendix).

## 5. Implications of the Mandate on the "Take-up Rate" by the Previously Uninsured

A major policy concern regarding the mandated Patient Protection and Affordable Care Act has been the degree to which it can succeed in achieving one of the central objectives of the mandate - inducing the uninsured to purchase health insurance. To reinforce compliance, the law imposes a sanction (or "tax") of $695 on those in the target population who choose to stay uninsured and avoid paying the premium by 2016. The ACA requires all tax payers to state on their annual tax reports whether they are enrolled in an accredited insurance plan. The IRS is charged with monitoring reporting and imposing the sanction upon discovery of non-compliance.

A few studies have used micro simulation models to assess various policy implications of the ACA (see, e.g., CBO, 2010). To our knowledge, however, none of these studies has taken into account the roles of self-insurance and self-protection in determining the "take-up", or compliance decision by the uninsured. Our calibrated simulations of the baseline and extended "full-insurance" model offer some direct insights about this issue.

### a. Experiment design

By incorporating into our baseline and extended model the relevant sanction imposed on non-switchers we can rerun our calibrated simulations to estimate the sanction's implications for the "full-insurance" decision, and thus the expected degree of compliance by the uninsured. The first issue we need to settle in this regard, however, is how to assess the *effective* sanction. This issue is relevant since in practice at least some of the uninsured will be able to avoid paying the sanction due to less than fully effective monitoring and enforcement procedures, or other evasive tactics, as is the case with all legal infractions. To deal with this issue, we apply 3 different scenarios regarding the effective penalty levels.

*Penalty level 1*: We take enforcement to be fully successful and impose a sanction of $695. Since the private share of the average employee health insurance premium reported by MEPS is $960, the fully enforced sanction would amount to 72.4% of the premium, which seems unrealistic.[16]

*Penalty Level 2*: We impose a sanction of $455 to achieve a 50% compliance rate, which is the rate experience by Massachusetts according to Census data (see Yelowitz and Cannon, 2010). At this level, the penalty amounts to 47.4% of the average employee share of the $960 premium.

*Penalty Level 3*: We impose a sanction of $222.4. This figure is the fraction of the mandated sanction of $695 to the actual premium of $3,000 to be charged for the "silver plan" offered by the ACA-established Central Exchanges, which is 23.2% = $695/$3000. This accounts for the premium costs incurred by those who purchase private insurance policies. Applying this rate to the average premium of $960 yields an effective penalty of $222.4.

To what extent would the alternative penalty levels assure compliance?

### b. Results:

*By our calibrated simulations of the baseline model*, when we account for 4 available insurance measures, including the 3 alternatives to market insurance – SI, SP, and the safety-net measure – the estimated compliance rates by the uninsured range from 24.5% of the target population for the lowest penalty to 75% for the highest, as shown in Panel A of the baseline-model in Table 7.

If SI, SP, and the Safety-net system are ignored, however, the compliance rates would be much higher, ranging from 57.7% for the lowest penalty to 85.2% for the highest penalty. By Panel B of section I of Table 7, the overstated compliance rates in Panel B relative to A would then range from 13.6% = [85.2%/75% - 1] to 135.7% = [57.75%/24.5% - 1].

*By our calibrated simulations of the expanded model*, if we account for all 4 available insurance measures, the compliance rates range from 26% for the lowest penalty to 76.5% for the highest. If the three alternatives to market insurance are ignored, the compliance rates are again much higher, ranging from 53.7% for the lowest penalty to 84.25% for the highest. The overstated compliance rates in Panel B relative to A then range from 10.1% to 106.5%, respectively.

Our estimates of the fractions of the target population remaining uninsured despite the expected sanctions are of the same order of magnitude in both models, but the compliance rates are generally higher in the extended model when all 4 insurance measures are accounted for and lower when the 3 alternatives to market insurance are ignored. The overstated compliance rates are thus lower in the extended model. According to both models, however, compliance rates that ignore the role of all forms of SISP could overstate the rates that recognize this role by over 50% on average, as illustrated by the average overstated compliance rates in part II of Table 7.

### c. Linking with CBO estimates

The CBO has not reported direct estimates of the "take-up" rate of the uninsured population. However, according to the CBO (2010) report assessing the effects of the insurance-coverage provisions of the Reconciliation Proposal, Combined with H.R. 3590 as passed by the Senate, 52 million nonelderly people would be uninsured under the current law in 2016. The report also sets the post-policy uninsured nonelderly at 21 million. This implies a compliance/take-up rate of (52-21)/52 = 59.6%, which is more in line with the forecasted compliance rates reported in Table 7 for the lowest penalty levels if one ignores the role of the private alternatives to market insurance we address in this paper. This indicates that estimates of the increase in the insured population, such as those derived by the CBO, may indeed be significantly overstated.[17]

## 6. The Insurance Mandate's Effects on Health Spending and Health Benefits under a uniform premium

Our analysis of the problem of the uninsured in the context of the full-insurance decision also offers some insights into the other major objectives of the insurance mandate – lowering the health care costs and improving the health status of the uninsured. It is arguable, of course, that viewing the decision to be uninsured as a voluntary choice cannot improve the individual welfare of the previously uninsured who would be induced by the force of the sanction to purchase health insurance they previously eschewed in favor of self-controlled health care. But the question remains about two objective welfare indicators: whether the mandated ACA would lower the overall health care system's *costs,* and improve its overall health *benefits.*

A. Cost-Benefit Analysis

The net effects are an open issue. The sanction would lower the costs of the safety-net system, but it would also increase spending on insured medical services. More important, it would bring about a net reduction in self-insurance and self-protection, which can produce adverse outcomes for per-capita health.

Both the baseline model and the extended model provide insights into this issue via calibrated simulations by solving for the full-insurance spending decision. But the baseline model has an important limitation in this regard because it does not recognize any benefits to health associated with market insurance other than the smoothing of income losses associated with ill-health. Indeed this model shows that those induced by the sanction to become insured would experience a net increase in both spending on insurance and the expected illness losses (net of spending on SISP) while those staying uninsured experience a very minor decrease in such losses.

The extended model, by contrast, can in principle account for the mandate's effect on the demand for all four components of the full-insurance and show the net impact on both health-care spending and health outcomes. To accomplish the task we have used the calibrated simulations of the extended model to derive the equilibrium solutions for both the pre-ACA and post-ACA systems and thereby compute the behavioral shifts induced by the mandate on self-insurance (c*), self-protection (s*), and safety net spending (SN), as well as on remedial medical care outlays M*– both insured and out of pocket – which are imputed using the production functions (4), (6) and (12).

B. Simulation design using the extended model

a. Computing the change in the full Costs of Care generated by the mandate

To arrive at an overall estimate of the cost of health care under our stylized versions of the ACA relative to the pre-ACA "free-market" system, we go through the following steps. We first produce the solution of the model for the entire sample population, stratified by the endowed probabilities of health hazards, $p_j^e$, which gives us the values of the levels of optimal spending on the four "insurance" components: c*, r*, $L^0$, and M* for any given endowed probability. Next, we impose the alternative magnitudes of the ACA's effective sanctions for the remaining uninsured and reproduce the model's behavioral solutions for the optimal spending on the four components of the full insurance decision at all population grids. We also recover the new critical value of $p_1^{*e}$ separating the population of insured from that of the uninsured. This enables us to compute separately the aggregate changes in optimal spending on all components of the full insurance choice, including medical care costs, by members of the previously uninsured population who are induced to become "switchers" (i.e., those with endowed probabilities higher than $p_1^{*e}$), relative to those who remain "non-switchers" (i.e., those with probabilities lower than $p_1^{*e}$). We then compute the per-capita values of the changes in spending and expected utility for the group of switchers and non-switches, i.e., the values for the average person in each group.

A special challenge exists in connection with estimating the overall levels and changes in the dollar values of the safety-net medical care costs for which no data are reported. Our simulated model does solve for the magnitude of the level of health benefits produced by the safety net, estimated as $L^0$ per-capita. We have therefore recovered the implicit levels and changes in the medical care inputs, $M^0$, used in the production of $L^0$ by inverting the production function $L^0 = L^a[1-\exp(-c_M M^0)]$ and then calculating the dollar input costs per-capita for both the switchers and non-switchers as the group's expected safety-net costs it imposes for the average person as $\int_0^{p_1^{*e}} B(r^*)p^e P_M M^0 dp^e$.

b. Computing the overall change in the health and welfare benefits generated by the mandate.

Having solved for the changes in the optimal values of c*+r*, M* and $M^0$ per-capita for the groups of switchers vs. non-switchers, as described in the previous section, we are now in a position to compute the net changes in expected health and welfare benefits, or alternatively the expected changes in health losses generated by the mandate per-capita. Note that the net effects of the imposed sanction on the total change in health care costs and benefits per-capita for each group incorporate the effects on four distinct sources of change: spending on SISP, the safety net, and medical care services (insured and out-of-pocket), as well as the change in the composition of the previously uninsured (switchers vs. non-switchers) induced by the sanction.

C. Simulation results

In the following analysis we illustrate the changes by assuming that the effective sanction for non-insurance is sanction level 2 - 47.3% of the premium - although the effects of other sanction levels are also illustrated. Table 8a reports the net effects of the shift on all health care outlays, while Table 8b reports the net effects on health and welfare outcomes, per capita.

*Impact on health care costs*: Table 8a shows the effects on the total health care outlays. For those becoming insured (the "switchers") the simulations show that the mandate induces a reduction in SISP spending per-capita of $27.12 or 14.62% relative to their level under the pre-ACA system (here assumed for convenience to involve the opportunity costs of strictly self-efforts such as diet and life-style changes) since the model shows the latter to be joint substitutes for market insurance. Even though the switchers no longer rely on free safety-net services, as we assume in the extended model, they now pay for all insured services in premiums they have previously deemed to be too high. The loss of spending power produces a negative "income effect" which accounts for part of the reduction in SISP according to our comparative statics analysis (see Appendix A2). In contrast, however, insurance lowers the effective price of the medical care services down to its co-payment level for the switchers, who previously paid for these services out of pocket. This would lead to "excess consumption" of medical care services per-capita of $290.69 (a 61.25% increase), which the literature ascribes to "ex-post moral hazard". In addition, we impute the change (savings) in the cost of medical care previously imposed by the uninsured switchers through their use of the safety-net system to amount to $56.37 per-capita. The net effect would be an increase of $207.20 in health care spending per-capita or 28.92%[18].

As for those who choose to stay uninsured, they would now have to bear the cost of the penalty, which also reduces their spending power. The negative "income effect" produced by this change would lower their spending on SISP by $2.27 per-capita (-1.64%) and their out-of-pocket spending on medical care services (M) by $3.06 per-capita (-1.86%), but would lead to an increase, albeit negligible in the use and cost of safety-net medical services. The net effect on the total health care costs induced by the change in behavior by this group thus amounts to a reduction of $5.31 per-capita, or 1.65% (see fn. 17).

The bottom line is that the net changes in the total health care costs of the mandated ACA, relative to the non-ACA system due to all the behavioral changes by the respective groups, thus amounts to a rise of $100.95 per capita, or a total of $30.6 billion for the economy as a whole, relative to the total costs of the pre-ACA system of $519.28 per capita, or a total of $157.19 billion for the economy – an increase of 19.44% (see fn.17).

We should note again that our estimates are based on our projections using the MEPS data from 2009. These estimates cannot be compared to the actual costs incurred by the ACA through 2016, because the latter reflect major subsidies provided by the Federal government to both individuals and insurance companies, partly financed by cutting benefits to Medicare, as well as cost reduction achieved by shifting many of the potential ACA subscribers into Medicaid.[19] The sum total of these hidden costs cannot be easily assessed. Several sources indicate that subsidies may have well exceeded the $716 billion cut from Medicare, if we also tally the hidden costs associated with the added spending on Medicaid.[20]

*Impact on health outcomes:* Table 8b illustrates the impacts of the behavioral changes induced by the mandated ACA system relative to the pre-ACA system on the both prospective health losses from illness and the ultimate health and welfare outcomes. These effects are shown to be asymmetrical across those who are induced to switch and those who choose to remain uninsured. The net effects on welfare, which also take into account the changes in consumption benefits induced by the mandated ACA, however, are predictably symmetrical.

For switchers, the reduction in SISP due to the substitution and income effects of paying the full premium they previously eschewed leads to higher prospective health losses, i.e., an expected reduction of 0.3036 units or 6.62% from the endowed health level per-capita. The significantly higher spending on medical care services (61.25% by table 8a), however, produces a net *positive* improvement in the health status of this group, imputed from equation (12), although the increase is quite small, just 0.48% under all sanction levels.

For non-switchers, the income effect of the health sanction unambiguously worsens their health level because of a decrease in both SI and SP, leading to an expected decline of 0.0018 health units, or a 0.11% fall in endowed health, as well as a reduction of 1.86% in the demand for out-of-pocket medical care services. The slight increase in use of safety-net services restores only a negligible amount of the expected health loss.

The magnitudes of the overall net effect on the health benefits enjoyed by both groups can in principle depend on the size of the imposed sanctions. The loss to non-switchers becomes more pronounced as the expected penalty level becomes higher (a critical level close enough to the insurance premium would actually reverse their decision to remain uninsured). But as Table 8a reveals, the favorable net effect on the switchers outweighs the unfavorable effect on those remaining uninsured, essentially because the switchers substantially increase their consumption of health care services.[21]

The higher overall health benefits, however, comes at a significant cost. Under all the penalty levels considered in Table 8, the percentage increase in expected outlays on medical care services far exceeds the percentage reduction in the health loses due to the improved health benefits. More generally, the induced spending on medical care, some of which may be inefficient (to the extent it represents over-consumption or ex-post moral hazard) come at the expense of ordinary consumption for all members of the previously uninsured group.

The bottom line is that the mandate produces an increase in the ultimate health level under the ACA relative to the pre-ACA system. But this increase in "recovered health" comes at a significant cost. Under all penalty levels considered in Table 8a, the percentage increase in expected outlays on medical care services, despite the savings in the safety net costs generated by the previously uninsured who become insured, far exceeds the percentage reduction in health loses, as table 8b indicates. More important, the increase in health levels due to the increased use of insured remedial care services comes at the expense of a decline in preventive health care services which are served by SISP – note that a small percentage decline in SISP yields a bigger increase in expected health losses, which exposes the individual to larger future health risks.

*Impact on expected utility*: More generally, the large increase in spending on medical care services by the switchers (representing over-utilization, or ex-post moral hazard induced by a low coinsurance rate) comes at the expense of a fall in ordinary consumption spending. This is the case also for the non-switchers who must now bear the cost of the sanction. The net effect on expected utility winds up being unfavorable for both the switchers and the non-switchers. This result follows predictably from our basic economic approach since we assume that all members of the previously uninsured group could have previously chosen to become insured but decided in favor of the no-insurance option.

## 7. Allowing for hypothetical premium changes

The last conclusion has an important caveat. It follows from our assumption that the price or premium for insurance remains the same under the model's stylized versions of the ACA and the pre-ACA systems. Indeed, the realization that the mandated sanction can be welfare reducing is perhaps one of the main reasons that the ACA has added a number of supplementary provisions that aim at subsidizing or lowering premiums for families in the target group with incomes up to 400% of the poverty line. These include placing caps on the allowable out-of-pocket premiums, providing subsidies for buying insurance, and inducing competitive premiums through state-based health insurance exchanges. Clearly, a large enough decrease in insurance premiums could in principle improve the health and welfare outcomes for a larger fraction of the previously uninsured, who would be influenced to switch by the carrot of lower premiums as well as by the stick of the mandated sanction for non-insurance.

There are also important reasons, however, to expect that the premiums under the ACA would actually wind up being higher than those in the pre-ACA system. The important contributing factors are the expanded coverage provisions of the ACA's "guaranteed issue", which require, e.g., that premiums would not be raised to reflect adverse pre-existing health conditions or gender differences in endowed health, and that family insurance plans would cover children up to age 26. In a competitive system, these regulations would significantly raise the premiums to cover the added costs to health insurance providers.

To gauge the sensitivity of Tables 8a and 8b's results to the assumption of a fixed premium, we conduct calibrated simulations which allow insurance premiums to change up and down by 10% beyond their current levels and re-estimate the effects of these changes on the relative magnitudes of the expected costs and benefits of an ACA-like mandate holding all other model parameters intact. The detailed results under penalty level 2 are reported in Tables 9a and 9b.

Not surprisingly, a 10% increase in premiums increases the costs of and spending on health care, and reduces the health and welfare benefits to both switchers and non-switchers; the converse occurs if premiums fall by 10% (see Tables 9a and 9b). The non-switchers, however, are hit harder by the increase in the premium, and benefit less from a reduction in premiums. The costs and benefits are shown to get larger if the effective penalty is higher.

One can address a related, and a more relevant question in this context: by how much would the premium need to fall to maintain the pre-ACA level of personal utility under the ACA system?

What we find is that for the 3 levels of penalties in table 8a, the required compensation for the switchers alone, (disregarding the non-switchers who cannot benefit from lower premiums) would entail lowering the premiums by 44.03%, 38.5%, and 20.5%, respectively. The compensating reduction in premiums for the whole group of the previously uninsured, measured as the premium reduction to the switchers that would overcome the welfare decline for the switchers would need to rise to 44.35%, 42.9%, and 38.85%, respectively. Such a fall in the premium level would bring about a large reduction in SISP, an increase in the prospective losses from illness (pL)*, and an explosion in remedial medical care spending (see Tables 9a and 9b).

## 8. Conclusion:

We can summarize our findings concerning the problem of the uninsured by returning to the basic questions we pose in the introduction: to what extent do the private alternatives to the conventional market health insurance account for the *magnitude* of the problem; how important are they in providing "insurance" services; and what policy implications can we draw concerning the effectiveness of health insurance reforms, such as the ACA, in insuring the uninsured and improving health care. We have attempted to answer these questions by developing and implementing a "full insurance" model that accounts for the interactions among the major components of the full-insurance system through calibrated comparative statics simulations.

Our answers are subject to a number of limitations. Our stylized model captures only the basic features of both the health insurance system and the ACA reform. We therefore abstract from much of the detailed aspects of the health insurance market and from the indirect effects the ACA may have, e.g., on the labor market and the prevailing premiums, although we do consider the potential effects of possible changes in insurance premiums on all outcomes. Yet our analysis has direct implications about the impact of the ACA on all the full-insurance choices. The consistency of the results we obtain from our calibrated simulations of both the baseline and extended models indicate that the qualitative and quantitative insights we gain are non-trivial.

The reason is both methodological and practical. Adding self-efforts as alternatives to the formal insurance market puts the "problem of the uninsured" in a context in which the offsetting behavioral interactions between MI and SISP are accounted for. This is especially relevant in the case of health insurance. Fast growing scientific evidence suggests that maintaining proper diet and exercise, avoiding hazardous consumption and risky life styles, and, more generally, pursuing what we call SISP, or "preventive care" measures, plays a critical role in avoiding or limiting health losses from illness, which is not less decisive than the role of insured "remedial care" services in preventing illness and restoring health after illness strikes.

Despite the simplifying assumptions, the calibrated simulations of both the baseline model and the extended model match quite well the empirical evidence from 2009 provided by MEPS about the percentage of the uninsured and the average medical expenditures on health care services targeted by simulations. Notwithstanding their differences, both models also reveal consistent patterns of quantitative solutions for the actual demand for, or spending on, SISP as well as their relative impact on the magnitude of prospective losses from illness.

Our calibrated simulations in Tables 1 and 4 indicate that self-insurance and self-protection account for 31.3% of the uninsured by the baseline model, or 28.2% by the extended model. Jointly with the safety net system, these alternatives account for 50.3% and 45.5% of the uninsured, respectively, in 2009. Tables 3 and 6 indicate that optimal efforts devoted to "protective" care, especially via SP but also via SI, lower significantly the "endowed" probability and severity of losses from ill health and their expected real costs by between 20-30%. Consistent with our analytical predictions, the calibrated simulations also indicate that under the given structure of health insurance plans, SI and SP are jointly substitutes for market insurance and for insured remedial medical care services. These results have important policy implications.

As our analysis in Section 5 and Table 7 illustrates, estimates of compliance rates with the mandated provision of the ACA by the previously uninsured could be significantly overstated if no account is given to the role that SISP and the safety net have played in motivating the original decision of individuals to be uninsured. The precise estimates depend on the magnitude of the penalty actually imposed on non-compliers, but our illustrated results indicate that it might be significantly overstated, perhaps by over 50% (also see fn. 17).

The analysis in Section 6 and Tables 8a and 8b indicate that although the mandated ACA sanction results in an overall improved health benefits for those who are induced by the force of the sanction to become insured (the switchers), these benefits come at a significant rise in the health care costs impose by the switchers, even after accounting for the savings they generate in the costs of the charitable and publicly financed costs of the safety net system. Our simulations in Table 8a indicate that these savings, as well as the lower spending on SISP by both switchers and non-switchers are more than offset, however, by increased spending on total medical care services by switchers, which rise by 28.92%, and the overall net increase in spending on health care outlays by both switchers and non-switchers is 19.44% under sanction level 2. Moreover, while the switchers wind up with a net gain in their health level, however, the non-switchers, who find it optimal to absorb the costs of the sanction and stay uninsured, are net losers. Tables 8a and 8b indicate that their health care costs rise, while their health level falls.

There is more to the tradeoff between the reduction in SISP by both switchers and non-switchers and the increased access to insured medical services gained by switchers. As modeled in our analysis, SISP provide protective health care benefits that reduce prospective health losses - the probability and severity of losses if illness strikes. Insured medical services, by contrast, essentially provide remedial care which helps restore or rebuild lost health when illness occurs. Although our model treats the benefits provided by these alternative services as additive, this may be valid in the short term. The greater susceptibility and exposure to illness due to reductions in preventive care, however, may be more detrimental in the longer term. In this context, the ACA provisions which mandate health policies to cover preventive as well as remedial medical services are a step in the right direction.

Furthermore, our calibrated simulations in Table 8b indicate that the improved health benefits to the switchers are more than offset by a larger reduction in their regular consumption benefits because of the premium costs they now need to bear. This is also the case for the non-switchers because of the fine they have to absorb. Both results are predictable by basic economic theory, since they result from a mandated sanction, rather than voluntary choices. These losses to both groups can in principle be ameliorated or even eliminated by lowering the effective premiums to the switchers. As our sensitivity analysis in Section 7 and tables 9a and 9b indicates, however, the compensating reductions in health insurance premium that would prevent a fall in expected utility to those who choose to switch may be as high as 38.5% (under sanction level 2).

There may be other benefits to the uninsured or society as a whole that would be generated by the ACA reform plan, which our models do not account for. For example, the models do not consider distortions in the private health insurance or health care markets which result in denial of access to the insurance or health care markets. Our calibrated simulations could also be made more accurate by more detailed stratifications of the target population affected by ACA. But our model and projections based on the data available in 2010 at the time of the passage of the ACA law have been generally consistent with the empirical evidence in the 4 years following the program's launching: the program has fallen short of the expected take up of insurance by the uninsured relative to what CBO had originally projected in 2010, based on comparable data. Also the actual costs of the program, including the federal subsidies offered to individuals and insurance companies through various government programs (see fn. 19) may have been consistent with our expectation for an actual rise by 19.44%. These factors may have also contributed to the degree of public dissatisfaction with the ACA as revealed by public opinion polls, and current plans to abolish or reform the system.

The central message of this paper is that useful analyses of the problem of the uninsured must recognize that a major factor accounting for eschewing insurance is the real price of insurance relative to the shadow prices of self-insurance and self-protection. As Proposition 2 and the calibrated numerical analysis indicate, it is the significant deviation of the real price of insurance faced by individuals relative to both the actuarially fair price and the shadow prices of self-insurance and self-protection, which is largely responsible for the problem of the uninsured, as well as for excess spending on insured medical services. Efficient reforms of the health insurance system should allow for the emergence of competitive health insurance premiums that limit the deviations of the price from its actuarially fair level, since this would work to maximize subscription to health insurance policies, while also promoting self-insurance and self-protective behaviors that enhance preventive care and the effectiveness of remedial care. This can be done through competitive market solutions that allow for the establishment of risk pools and associated premium structures that better reflect individual efforts at self-insurance and self-protection, and offer individualized coverage plans that account for the shadow prices of the health care services offered for remedial care. Such reforms could contribute to the cost effectiveness of health insurance market, as well as bolster the health benefits of the overall ‘full insurance’ system (see Ehrlich and Yin, 2013).

## References


Agency for Healthcare Research and Quality (2009), Medical Expenditure Panel Survey, http://meps.ahrq.gov/mepsweb/.

Brady, Kevin (2017), "Improving Access to Quality, Affordable Health Care for Low Income Americans", https://waysandmeans.house.gov/wp-content/uploads/2017/03/ImprovingAccesss_AHCA.pdf

Briani, Marco, Magnani, Marco and Mario Menegati (2017), "Optimal choice of prevention and cure under uncertainty on disease effect and cure effectiveness", *Research in Economics",* in press.

Chang, Yang-Ming, and Isaac Ehrlich (1985), "Insurance, Protection from Risk, and Risk Bearing", *Canadian Journal of Economics*, 18 (3), 579-86.

Congressional Budget Office (2010), Report on H.R. 4872 the Reconciliation Act of 2010. http://www.cbo.gov/publication/21327.

Congressional Budget Office (2012), Payments of Penalties for Being Uninsured Under the Affordable Care Act, September 19, 2012. http://www.cbo.gov/publication/43628.

Congressional Budget Office (2016), Federal Subsidies for Health Insurance Coverage for People Under Age 65: 2016 to 2026, updated at https://www.cbo.gov/publication/51385

Ehrlich, Isaac, and Gary Becker (1972), "Market Insurance, Self-Insurance and Self-Protection", *Journal of Political Economy*, 80 (4), 623-48.

Ehrlich Isaac, and Yong Yin (2013), "Equilibrium Health Spending and Population Aging in a Model of Endogenous Growth: Will the GDP Share of Health Spending Keep Rising?", *Journal of Human Capital,* 7(4), 411-447.

Jost, Timothy (2017), "Insurers Seeking Risk Corridor Payments Get First Courtroom Win", HealthAffairs Blog, http://healthaffairs.org/blog/2017/02/10/insurers-seeking-risk-corridor-payments-get-first-courtroom-win-irs-updates-qas/.

Kunreuther, Howard, and Adam Z. Rose eds., *The Economics of Natural Hazards*, The International Library of Critical Writings in Economics, Mark Blaug, Series Editor, Edward Elgar Publishing Ltd., U.K, 2004.

Manning, W. G. and M. S. Marquis (1996), "Health Insurance: The Tradeoff Between Risk Pooling and Moral Hazard", *Journal of Health Economics*, 15(5), 609-639.

Yelowitz Aaron and Michael F. Cannon (2010), "The Massachusetts Health Plan – Much Pain, Little Gain", *Policy Analysis*, No. 657, the Cato Institute.

**Table 1: Effects of Alternative Feasible "Insurance" Option on % Uninsured in Target Population: Upper Limits (Baseline Model)**

| Feasible "insurance" options | % uninsured |
|---|---|
| All options (4) | 20.00 |
| Market Insurance, Self-Insurance & Self-Protection (no safety net) | 16.20 |
| Just Market Insurance & Safety-net | 13.75 |
| Just Market Insurance Only | 11.11 |

The calculation is based on following set and calibrated parameters in the baseline model:

$\sigma = 2$; $I = \$36{,}000$; $L^e = \$8{,}250$; $L^a = (.8)\, L^e$; $R = \$960$; $L^0 = 12.73\%\, L^e$

$A(c) = A_h + (1-A_h)\exp(-\eta_1 c)$ with $A_h = 0.8$ and $\eta_1 = 0.05$;

$B(r) = B_h + (1-B_h)\exp(-\eta_2 r)$ with $B_h = 0.7$ and $\eta_2 = 0.05$

**Table 2: Optimal Spending on Self-Insurance and Self-Protection: Baseline Model**

| Endowed probability of sickness | If Insured* | | | If Uninsured* | | |
|---|---|---|---|---|---|---|
| | c (in dollars) | r (in dollars) | c+r as a % of premium | c (in dollars) | R (in dollars) | c+r as a % of premium |
| 0.1 | 38.26 | 20.01 | 6.07 | 42.29 | 45.98 | 9.19 |
| 0.2 | 51.99 | 19.95 | 7.49 | 55.10 | 58.91 | 11.88 |
| 0.3 | 60.05 | 19.94 | 8.33 | 62.51 | 66.36 | 13.42 |
| 0.4 | 65.78 | 19.93 | 8.93 | 67.66 | 71.54 | 14.50 |
| 0.5 | 70.23 | 19.92 | 9.39 | 71.57 | 75.46 | 15.32 |

* At $p^e = 0.1$ and $0.2$ consumers are optimally uninsured. At higher values of $p^e$ they are optimally insured.

See notes to Table 1 for calibrated parameters.

**Table 3: Impact of Optimal Self-Insurance and Self-Protection on Prospective Sickness Losses (Baseline Model)**

| Endowed probability of sickness | If Insured | | | If Uninsured | | |
|---|---|---|---|---|---|---|
| | $L^*/L^e$ in % | $p^*/p^e$ in % | EL* in $ | $L^*/L^e$ in % | $p^*/p^e$ in % | EL* in $ |
| 0.1 | 82.95 | 81.03 | 5,545 | 82.41 | 73.01 | 4,964 |
| 0.2 | 81.49 | 81.06 | 5,449 | 81.27 | 71.58 | 4,799 |
| 0.3 | 80.99 | 81.07 | 5,417 | 80.88 | 71.09 | 4,743 |
| 0.4 | 80.75 | 81.08 | 5,401 | 80.68 | 70.84 | 4,715 |
| 0.5 | 80.60 | 81.08 | 5,391 | 80.56 | 70.69 | 4,698 |

Note: EL*, expected sickness losses, is defined as $EL^* = p^*L^*$, where $L^* = A(c^*)L^e$ and $p^* = B(r^*)p^e$.

See note to Table 1 for calibrated parameters

**Table 4: Effects of Alternative Feasible "Insurance" Option on % Uninsured in Target Population: Upper Limits (Extended Model)**

| Feasible "insurance" and "protection" options | % uninsured |
|---|---|
| All options (4) | 20.00 |
| Market Insurance, Self-Insurance & Self-Protection (no safety net) | 16.49 |
| Just Market Insurance & Safety-net | 14.36 |
| Just Market Insurance Only | 11.87 |

The calculation is based on the following set and calibrated parameters for the extended model:

$\sigma = 2$; $I = \$36,000$; $R = \$960$; $\kappa = 25\%$; $Pm = 1.75$

$H^e = 11,300$; $L^e = 0.005H^e$; $L^a = (.8)\,L^e$; $L^0 = 0.336L^e$;

$\theta = -7$; $\eta_3 = 0.00183$

$A(c) = A_h + (1-A_h)\exp(-\eta_1 c)$ with $A_h = 0.8$ and $\eta_1 = 0.05$;

$B(r) = B_h + (1-B_h)\exp(-\eta_2 r)$ with $B_h = 0.7$ and $\eta_2 = 0.05$

**Table 5: Optimal Spending on Self-Insurance, Self-Protection, and Expected Medical Expenditure (Extended Model)**

| Endowed probability of sickness | If Insured* | | | | If Uninsured* | | | |
|---|---|---|---|---|---|---|---|---|
| | EM** (in $) | C (in $) | r (in $) | c+r % of premium | EM** (in $) | c (in $) | r (in $) | c+r % of premium |
| 0.1 | 525 | 119.82 | 24.52 | 15.04 | 318 | 122.23 | 49.53 | 17.89 |
| 0.2 | 1,000 | 134.23 | 35.13 | 17.64 | 628 | 137.31 | 58.70 | 20.42 |
| 0.3 | 1,474 | 143.68 | 40.84 | 19.22 | 937 | 146.99 | 63.10 | 21.88 |
| 0.4 | 1,947 | 151.13 | 44.55 | 20.38 | 1,246 | 154.54 | 65.74 | 22.95 |
| 0.5 | 2,420 | 157.56 | 47.20 | 21.33 | 1,555 | 161.02 | 67.52 | 23.81 |

* At $p^e$ = 0.1 and 0.2 consumers are optimally uninsured. At higher values of $p^e$ they are optimally insured.

** EM is defined as expected medical expenditure as $B(r^*)p^ePmM^*$.

See note to Table 4 for calibrated parameters.

**Table 6: Impact of Optimal Self-Insurance and Self-Protection on Prospective Sickness Losses (Extended Model)**

| Endowed probability of sickness | If Insured | | | If Uninsured | | |
|---|---|---|---|---|---|---|
| | $L^*/L^e$ in % | $p^*/p^e$ in % | EL* | $L^*/L^e$ in % | $p^*/p^e$ in % | EL* |
| 0.1 | 80.05 | 78.81 | 35.64 | 80.04 | 72.52 | 32.80 |
| 0.2 | 80.02 | 75.18 | 33.99 | 80.02 | 71.59 | 32.37 |
| 0.3 | 80.02 | 73.89 | 33.41 | 80.01 | 71.28 | 32.22 |
| 0.4 | 80.01 | 73.23 | 33.11 | 80.01 | 71.12 | 32.15 |
| 0.5 | 80.01 | 72.83 | 32.92 | 80.01 | 71.03 | 32.11 |

Note: EL*, expected sickness loss, is defined as $EL^* = p^*L^*$, where $L^* = A(c^*)L^e$ and $p^* = B(r^*)p^e$.

See note to Table 4 for calibrated parameters

**Table 7: Assessing Compliance**

| Penalty Level (as % of premium) | % pop remaining uninsured | Change in % uninsured* | Compliance rate** | % Overstated compliance (B/A)# |
|---|---|---|---|---|
| **I. Baseline Model** | | | | |
| A. (Accounting for all 4 measures of "full insurance") | | | | |
| 1 (72.4%) | 5% | 15% | 75% | |
| 2 (47.4%) | 10% | 10% | 50% | |
| 3 (23.2%) | 15.1% | 4.9% | 24.5% | |
| B. (Ignoring the SI, SP, and safety-net alternatives ) | | | | |
| 1 (72.4%) | 2.96% | 17.04% | 85.2% | 13.6% |
| 2 (47.4%) | 5.72% | 14.28% | 71.4% | 42.8% |
| 3 (23.2%) | 8.45% | 11.55% | 57.75% | 135.7% |
| **II. Extended Model** | | | | |
| A. (Accounting for all 4 measures of "full insurance") | | | | |
| 1 (72.4%) | 4.7% | 15.3% | 76.5% | |
| 2 (45.8%) | 10% | 10% | 50% | |
| 3 (23.2%) | 14.8% | 5.2% | 26% | |
| B. (Ignoring the SI, SP, and safety-net alternatives) | | | | |
| 1 (72.4%) | 3.15% | 16.85% | 84.25% | 10.1% |
| 2 (45.8%) | 6.36% | 13.64% | 68.2% | 36.4% |
| 3 (23.2%) | 9.26% | 10.74% | 53.7% | 106.5% |

* Change in % uninsured is calculated as the difference between the initial % uninsured (20%) and remaining % uninsured under penalty.

** Compliance rate is calculated as ratio of the change in % uninsured and the initial % uninsured.

# Overstatement % if assessment ignores the existence of SIST and safety net (NS) alternatives

**Table 8a: Changes in Health Cost Per-Capita under the Mandate at Alternative Effective Penalty Levels (Extended Model)**

| | SISP ($) | | | Safety-Net Care ($) | | | Expected Medical Care Spending[1] ($) | | | Overall net change[2] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Pre-ACA | Change | % change | Pre-ACA | Change | % change | Pre-ACA | Change | % change | Change | % change |
| Penalty Level 1: Penalty as 72.4% of premium[3] | | | | | | | | | | | |
| Switchers | 176.72 | -27.15 | -15.36% | 46.61 | -46.61 | -100% | 392.54 | 246.21 | 62.72% | 172.45 | 28.00% |
| Non-Switchers | 113.48 | -3.55 | -3.13% | 9.68 | 0.03 | -3.13% | 81.68 | -2.32 | -2.84% | -5.84 | -2.85% |
| Combined[4] | 161.86 | -21.60 | -13.34% | 37.93 | -35.65 | -93.99% | 319.49 | 187.81 | 58.78% | 130.55[4] | 25.14% |
| Penalty Level 2: Penalty as 47.4% of premium[3] | | | | | | | | | | | |
| Switchers | 185.45 | -27.12 | -14.62% | 56.37 | -56.37 | -100% | 474.60 | 290.69 | 61.25% | 207.20 | 28.92% |
| Non-Switchers | 138.27 | -2.27 | -1.64% | 19.49 | 0.02 | 0.10% | 164.37 | -3.06 | -1.86% | -5.31 | -1.65% |
| Combined[4] | 161.86 | -14.70 | -9.08% | 37.93 | -28.18 | -74.29% | 319.49 | 143.82 | 45.02% | 100.95[4] | 19.44% |
| Penalty Level 3: Penalty as 23.2% of premium[3] | | | | | | | | | | | |
| Switchers | 191.15 | -26.92 | -14.08% | 65.20 | -65.20 | -100% | 548.87 | 330.55 | 60.22% | 238.43 | 29.61% |
| Non-Switchers | 151.57 | -1.15 | -0.76% | 28.35 | 0.01 | 0.04% | 238.89 | -2.28 | -0.95% | -3.42 | -0.82% |
| Combined[4] | 161.86 | -7.85 | -4.85% | 37.93 | -16.90 | -44.56% | 319.49 | 84.26 | 26.37% | 59.46[4] | 11.45% |

[1] Expected health expenditure per capita is calculated as $\int_0^{p_1^{*e}} B(r^*) p^e P_M M^0 dp^e$ .

[2] The net effect on costs is calculated as the sum of changes in SISP, safety-net care, and expected medical care spending. The percentage changes in overall totals for each group, as well as for the combined group, are discussed in the text.

[3] The employee share of the premium is calibrated as $960.

[4] Overall combined changes are calculated as a weighted average of switchers and non-switchers, with weights determined by the groups' shares in the population of the uninsured. The latter are optimally determined at alternative effective penalties

**Table 8b: Changes in Health Level Per-Capita under the Mandate at Alternative Effective Penalty Levels (Extended Model)**

| Group | Change in expected loss from sickness[1] | | Change in expected health status[2] | | Change in expected utility (x $10^{-10}$) |
|---|---|---|---|---|---|
| | Change | % of expected loss from sickness | Change | % of expected loss from sickness | |
| Penalty Level 1: Penalty as 72.4% of premium[3] | | | | | |
| Switchers | 0.2929 | 7.24% | 0.01936 | 0.48% | -3.04 |
| Non-Switchers | 0.0025 | 0.30% | -0.00108 | -0.17% | -5.88 |
| Penalty Level 2: Penalty as 47.4% of premium[3] | | | | | |
| Switchers | 0.3063 | 6.62% | 0.02358 | 0.48% | -1.99 |
| Non-Switchers | 0.0018 | 0.11% | -0.00114 | -0.08% | -3.76 |
| Penalty Level 3: Penalty as 23.2% of premium[3] | | | | | |
| Switchers | 0.3157 | 5.58% | 0.02741 | 0.48% | -1.03 |
| Non-Switchers | 0.00098 | 0.04% | -0.00077 | -0.031% | -1.91 |

[1] Expected loss from sickness is calculated as $B(r^*)A(c^*)L^e$.

[2] Expected health is calculated as $EH = [1-B(r^*)]H^e + B(r^*)[H^e - A(c^*)L^e + L^a\varphi(M^*)]$ for the insured and $EH = [1-B(r^*)]H^e + B(r^*)[H^e - A(c^*)L^e + L^0 + (L^a-L^0)\varphi(M^*)]$ for the uninsured.

[3] The employee share of the premium is calibrated as $960.

**Table 9a: Changes in Health Cost Per-Capita Under the Mandate at Penalty Level 2**
**With Varying Premium levels: Sensitivity Analysis (Extended Model)**

| | SISP ($) | | | Safety-Net Care ($) | | | Expected Medical Care Spending ($) | | | Overall net change | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | Pre-ACA | Change | % change | Pre-ACA | Change | % change | Pre-ACA | Change | % change | Change | % change |
| 10% increase in premium | | | | | | | | | | | |
| Switchers | 188.12 | -27.53 | -14.63% | 60.23 | -60.23 | -100% | 507.10 | 305.75 | 60.30% | 218.02 | 28.86% |
| Non-Switchers | 144.72 | -2.28 | -1.58% | 23.37 | 0.02 | 0.09% | 196.99 | -3.70 | -1.88% | -5.96 | -1.63% |
| Combined | 161.86 | -12.25 | -7.57% | 37.93 | -23.78 | -62.69% | 319.49 | 118.53 | 37.10% | 82.51 | 15.89% |
| Original premium of $960 | | | | | | | | | | | |
| Switchers | 185.45 | -27.12 | -14.62% | 56.37 | -56.37 | -100% | 474.60 | 290.69 | 61.25% | 207.20 | 28.92% |
| Non-Switchers | 138.27 | -2.27 | -1.64% | 19.49 | 0.02 | 0.10% | 164.37 | -3.06 | -1.86% | -5.31 | -1.65% |
| Combined | 161.86 | -14.70 | -9.08% | 37.93 | -28.18 | -74.29% | 319.49 | 143.82 | 45.02% | 100.95 | 19.44% |
| 10% reduction in premium | | | | | | | | | | | |
| Switchers | 182.41 | -26.69 | -14.63% | 52.50 | -52.50 | -100% | 442.10 | 275.22 | 62.25% | 196.03 | 28.96% |
| Non-Switchers | 130.38 | -2.26 | -1.73% | 15.61 | 0.02 | 0.13% | 131.69 | -2.43 | -1.85% | -4.67 | -1.68% |
| Combined | 161.86 | -17.04 | -10.53% | 37.93 | -31.75 | -83.71% | 319.49 | 165.55 | 51.82% | 116.75 | 22.48% |

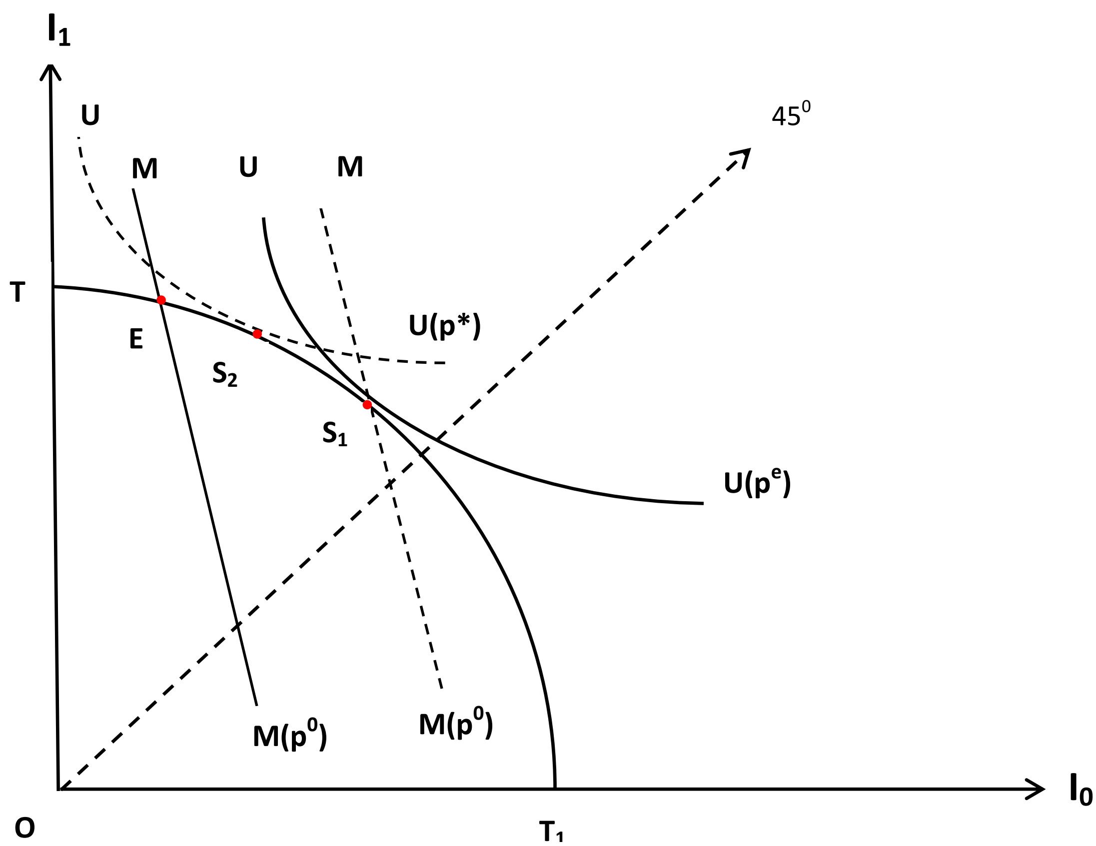
See notes to table 8a.

**Table 9b: Changes in Health Level Per-Capita under the Mandate at Penalty level 2**
**With Varying Premium Levels: Sensitivity Analysis (Extended Model)**

| Group | Change in expected loss from sickness[1] | | Change in expected health status[2] | | Change in expected utility (x $10^{-10}$) |
|---|---|---|---|---|---|
| | Change | % of initial expected loss from sickness | Change | % of expected loss from sickness | |
| 10% increase in premium | | | | | |
| Switchers | 0.3118 | 5.96% | 0.02509 | 0.48% | -2.55 |
| Non-Switchers | 0.0018 | 0.094% | -0.00132 | -0.07% | -3.81 |
| Original premium of $960 | | | | | |
| Switchers | 0.3063 | 6.62% | 0.02358 | 0.48% | -1.99 |
| Non-Switchers | 0.0018 | 0.11% | -0.00114 | -0.08% | -3.76 |
| 10% reduction in premium | | | | | |
| Switchers | 0.3004 | 6.59% | 0.02205 | 0.48% | -1.46 |
| Non-Switchers | 0.0017 | 0.13% | -0.00095 | -0.07% | -3.70 |

See notes to Table 8b.

**Figure 1: Rationalizing No-Insurance**



**Appendix: Estimating the role of the major determinants of the full insurance decision via comparative statics simulations**

**A1. The Baseline Model**

Table 2 summarizes the qualitative effects of shifts in the main parameters of the model on two main aspects of the full insurance decision: a. whether to insure or not to insure; and b. the impact of the choice on components of the full insurance choice, conditional on deciding to either i. insure, or ii. not to insure. The effects here apply just to optimal self-insurance and self-protection (SISP).

*Wealth effects*: As pointed out in Ehrlich and Becker (1972), the impact of an upward shift in "wealth", $W = I_1^e + \pi^e I_0^e$, depends on how the different endowments change. In reality, a higher endowed wealth has a conventional "income effect" but also increases the potential monetary loss, as wealthier individuals have higher opportunity costs of sick time. In Table 1 we thus show the impact of a "neutral, i.e., equi-proportional increase in the income endowments $I_1^e$ and $I_0^e$. Given our indemnity insurance structure and the CRRA utility function, the qualitative effects are driven essentially by the increase in exposure to loss, which unambiguously lowers the likelihood that the individual would choose[s] to remain uninsured, while also raising optimal spending on SI (c*) and SP (r*).

*The Endowed loss from sickness*: the effects are identical to those of the "neutral" wealth effect.

*The indemnity size*: Provides an incentive to insure, but the results for optimal SISP by the insured are ambiguous: since the change amounts to a higher endowment in state 0, the latter reduces the incentive to self-insure, but the effect may be ambiguous on self-protection. As the indemnity rises, the potential exposure to risk (effective size of loss) diminishes, but the income effect could be positive, at least initially. The uninsured are unaffected.

*The premium size*: The gross insurance price effect lowers the incentive to insure and increases the incentive not to insure, but it raises unambiguously the demand for SISP by the insured, as the latter are jointly substitutes for market insurance. The uninsured are naturally unaffected.

*Technological parameters affecting SI, SP*: Lower values of $A_h$ and $B_h$ synthesize two possibly conflicting effects: A partial decrease in each raises the marginal productivity of c and r in reducing A(c) and B(r). For example, a lower $A_h$ would thus result in a higher marginal productivity of c in effecting loss reduction, $A(c)L^e$. But $-\partial A(c)/\partial A_h = -[1-\exp(-\eta_1 c)] < 0$ exerts an independent scale effect, by lowering the *level* of A(c). The effects on A(c) or B(r) would then depend on the elasticity of each with respect to their respective arguments. The results are ambiguous also because SI and SP act like substitutes in this model.

*The risk tolerance parameter*: a larger σ reflects a higher tolerance for risk. Its impacts on the full-insurance components are generally predictable: a lower tendency to insure; but the effects on c* and r* can go in either direction as a result.

**A2. The extended model**

Table 2 summarizes the qualitative effects of shifts in the main parameters of the model on the full insurance decision by its three choices: a. to insure or not to insure; and b. their impact on components of the full insurance decision conditional on choosing to i. insure; or ii. not to insure. The effects here include optimal medical care spending as well as SIST.

*Income effects of parameter changes*

*Higher income* level: lowers the percentage uninsured and has positive impacts on the optimal SISP by both the insured and uninsured. Note that in the extended model, an upward shift just in $I_1^e$, hence effective income, W, as well, generates an income effect on the demand for health, H, and thus on the derived-demand for insured medical care inputs, M. The derived-demand for both c* and r* can also increase because of the health benefits they confer by lowering L*.

*Endowed loss from sickness:* has an ambiguous impact on the % uninsured because of a technical reason. If we do not allow $L^a$ to change proportionally, the impact is negative; otherwise the impact becomes positive. Yet, the impact on both SI and SP is always positive, as is the case in the baseline model, since the $L^a$ constraint does not limit the SISP's effectiveness.

*The size of maximum coverage*: reduces the percentage of uninsured, because a higher maximum coverage makes market insurance more attractive. The effects on SI and SP are ambiguous because of the interaction among all three measures of insurance.

*The premium size*: provides an unambiguous negative effect on the incentive to insure, or positive on the incentive not to insure, but an unambiguous positive effect on the demand for SISP by the insured, as the latter are substitutes to market insurance.

*The technological parameters affecting SI, SP* generate outcomes similar to those obtained in the baseline model.

*The risk tolerance and the degree of substitutability of H and M in consumption*: The σ effects are similar to those in Table A1. A higher θ indicates that M and X are more substitutable, which lowers the propensity to insure. For both the insured and the uninsured this lowers both c* and M*. The effect of θ on self-protection, however, is generally ambiguous.

*The Impact of basic parameters on medical Expenditures*: The impacts of upward shifts in basic parameters on medical care services, M, generally go in the same direction as those on self-insurance, except for the shifts in the level of safety-net services, $L^0$, and the relative risk aversion coefficient σ because of the interactions among the substitutable insurance measures.

**Table A1: Comparative Statistics (Baseline Model)**

| | | If Insured* | | If Uninsured** | |
|---|---|---|---|---|---|
| Parameter | % uninsured | c | r | C | r |
| "Wealth" | - | + | + | + | + |
| Loss from sickness, $L^e$ | - | + | + | + | + |
| Amount of Indemnity, $L^a$ | - | +/- | -/+ | **0** | **0** |
| Premium, R | + | + | + | **0** | **0** |
| Safety-net, $L^0$ | + | **0** | **0** | - | - |
| Σ? | + | +/- | -/+ | + | - |
| Ah | + | - | -/+ | + | - |
| Bh | - | + | + | + | - |

See table 1 for model parameters.

* Evaluated for $p^e = 0.24$.

** Evaluated for $p^e = 0.17$.

**Table A2: Comparative Statistics (Extended Model)**

| | | If Insured* | | | If Uninsured** | | |
|---|---|---|---|---|---|---|---|
| Parameter | % uninsured | M | c | r | M | c | r |
| Income, $I_1^e$ | - | + | + | + | + | + | + |
| Loss from sickness, $L^e$ | -*** | + | + | + | + | + | + |
| Maximum Coverage, $L^a$ | - | + | + | - | + | + | - |
| Premium, R | + | - | - | - | **0** | **0** | **0** |
| Safety-net, $L^0$ | + | **0** | **0** | **0** | - | + | - |
| σ | + | - | + | + | - | + | + |
| Ah | + | - | - | + | - | - | + |
| Bh | - | + | + | - | + | + | - |
| θ | + | - | - | +/- | - | - | -/+ |

See table 5 for model parameters.

* Evaluated for $p^e = 0.24$.

** Evaluated for $p^e = 0.17$.

*** An increase in $L^e$ raises the percentage of uninsured only since $L^a$ is fixed; accompanied by an equal proportion increase in $L^a$, however, a higher $L^e$ lowers the percentage uninsured.

## Endnotes

[1] The target population thus excludes the elderly, who are covered by Medicare, the very poor, who are covered by Medicaid, and uninsured children who are covered by the Children Health Insurance Program.

[2] The individual mandate was later declared to be a tax in the Supreme Court majority decision upholding the ACA law on June18 2012.

[3] See the 2010 study by LIMRA in "Rethinking Your Stance on Earthquake Coverage" by Liz Pulliam Weston, Los Angeles Times

[4] See "Households with life insurance hits lowest level in 50 years" by Sandra Block, 12.03.2010, USA Today

[5] The relevance of self-protection for preventive care has been discussed in a number of papers following Ehrlich and Becker's (1972) approach. For recent applications in health economics see Brianti *et al.* (2017) and related references mentioned therein.

[6] We generally assume that the utility function is strictly concave, U''< 0 at all levels of income in states 1 and 0, although for all the propositions in the following analysis we only need to impose the weaker assumption that the indifference curve system is downward sloping and convex toward the origin.

[7] The range of the transformation curve $TT_1$ assumes, for ease of illustration, that income transfers are not limited by the size of the endowed loss or even the possibility of gambling activity involving a reduction in income in the bad state 0 in return for a higher income in the good state 1.

[8] The first-order condition for an optimal value of self-protection requires that the shadow price (marginal cost) of self-protection (-1/p'(r)) – or its equivalent as specified in equation (10) – would be equal to the monetary value of the difference in the utility of net income in state 1 relative to state 0 (see Ehrlich and Becker 1972, eq. 28).

[9] Note that this condition applies regardless of whether individuals choose to insure or not to insure, because the assumed fixed price of insurance is not responsive to individual self-protection.

[10] This simplifying assumption is intended to fit the basic structure of the ACA reform. Upon further stratifying the population, however, we could also allow R to be age dependent.

[11] The exponential and convex form of the production functions is assumed for convenience since it sets a limit on the outcomes of both self-protection and self-insurance in fully eliminating the adverse physical and emotional consequences of major health losses even if c and r go to infinity. In a related paper about the role of individual health investments in the context of a general equilibrium growth model (see Ehrlich and Yin, 2013), similar convex production functions are assumed since they are essential for establishing a balanced growth equilibrium.

[12] Alternatively, we can solve for the percentage uninsured when SI, SP & safety net are not available. In this case, we find that 55.6% (11.11%/20%) of the target population will stay uninsured. But this is an absolute upper limit, since SISP are always available by Proposition 1 in Section 2. Note that the estimates in Tables 1 through 9b have been derived from projected data prior to the implementation of the effective ACA law in March 2010.

[13] Even if health is ultimately fully restored, the affected individual suffers a loss of good health benefits over the recovery period.

[14] Alternatively, we can solve for the percentage uninsured when SI, SP & safety net are not available. In this case, we find that 59.35% (11.87%/20%) of the target population will stay uninsured. But this is an overstated percentage, since SISP are always available to individuals and likely to be pursued by Proposition 1 in Section 2.

[15] The substitution relation between self-insurance and self-protection is verified in the extended model as well, in conformity with Proposition 3 in Section 2. We calculate that the optimal spending on c falls by 5.6%, from $137.05 to $129.38, for the uninsured (at an endowed loss probability of 0.14) when self-protection is made available. For the insured (at an endowed loss probability of 0.24) the change is from $145.93 to $138.36, representing a 5.2% reduction.

[16] Note that the estimate is excessive not just because of imperfect monitoring and enforcement, but also because the more attractive alternative would be to pay the premium and enjoy some insurance benefits.

[17] On September 19, 2012, the CBO published new estimates indicating that the population of the uninsured by current law would be 56 million in 2016, but will drop to only 30 million by 2016, indicating a roughly estimated compliance/take-up rate of 46.4%. Since CBO does not indicate that it has taken any account of the role of SISP and the safety net, our calibrated simulations indicate that this compliance estimate is still potentially overstated by magnitudes closer to the one we illustrate for penalty level 3 in part B of Table 7. Note, however, that the CBO projections are based on a target population that differs from our MEPS-based sample, however, as it includes unauthorized immigrants and people eligible for, but not enrolled in, Medicaid.

[18] 28.92% is the change in spending by switchers over its pre-ACA value 207.2/ 716.42, where $716.42 = 185.45 (SISP)+56.37 (SN)+ 474.6 (M*). Likewise, 1.65% is the change in spending by the non-switchers relative to its pre-ACA value 5.31/322.13, where $322.13 = 138.27 (SISP)+164.37 (M*)+19.49( SN). Pre-ACA total spending is thus 0.5*$716.42 + 0.5*$322.14=$519.28.

[19] For the same reason our projections of the take-up rate of the previously uninsured into the ACA cannot be compared with the actual rate, since the latter may have been significantly inflated by the shift of many in the target population into Medicaid.

[20] The original ACA has included three sources of subsidies designed to make the program affordable: offering direct subsidies and tax preferences for most Americans under age 65 through a variety of federal programs and tax preferences; providing subsidies to participating insurance company via the risk-corridor program, to help offset insurer losses in the first three years of the insurance exchanges; and shifting healthy people in the target group into the Medicaid programs (see Brady, 2017). A recent CBO report (CBO, 2016 updated), states that the federal government subsidies to individuals will total $660 billion in 2016. Jost (2017), reports the cost of the risk corridor program in 2015 was $5.8 billion. More recently, the impact of these indirect subsidies has been exposed through sharp rises in the average ACA premiums across most US states, caused partly by Congressional efforts to halt the payment of subsidies to insurance companies via the “risk corridor”. No official estimates are available for the value of the hidden subsidies provided through Medicaid.

[21] In this analysis we do not include any implications for the previously insured, essentially because it is not clear how their premiums are affected. See also our analysis in section 7.