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

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
JEL No. G22,H42,I13,I28

ABSTRACT

The problem of the uninsured – those eschewing the purchase of health insurance policies – cannot be fully understood without considering informal alternatives to market insurance called "self-insurance" and "self-protection", including the publicly and charitably-financed safety-net health care system. This paper tackles the problem of the uninsured by formulating a "full-insurance" paradigm that includes all 4 measures of insurance as interacting components, and analyzing their interdependencies. We apply both a baseline and extended versions of the model through calibrated simulations to estimate the degree to which these non-market alternatives can account for the fraction of the non-elderly adults who are uninsured, and estimate their behavioral and policy ramifications. Our results indicate that policy analyses that do not consider the role of self-efforts to avoid health losses can grossly distort the success of the ACA mandate to insure the uninsured and to improve the health and welfare outcomes of the previously uninsured.

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

Missing from the current policy debate regarding the desirability of the Patient Protection and Affordable Care Act of 2009 (ACA) is that being uninsured can be a rational choice even for people who are risk averse, and is not just an outcome of unaffordability. 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.[1] But there are other types of insurance against loss of life and property in which the voluntary purchase of insurance is far from universal. 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.[2] By comparison, less than 20% of the non-elderly population is uninsured. But only the latter case has generated an intense public debate about the need to make it mandatory for all people.

As is generally the case when individuals refrain from "entering" specific markets for goods and services, the explanation is that the market price sufficiently exceeds the individual consumer's optimal "entry price". In the case of market insurance against specific hazards, the analogous argument is that the price of insurance is sufficiently actuarially unfair for those for whom the probability of being at risk is relatively low, or because the insurance loading factor is high.

The main argument of this paper is that eschewing insurance is even more likely because of the existence of individually controlled alternatives to market insurance which have been termed "self-insurance" and "self-protection" (Ehrlich and Becker 1972). Individual self-insurance refers to actions people take to reduce their potential loss from the occurrence of specific hazards. Personal self-protection refers to actions individuals take to reduce the probability of loss occurring in the first place. These alternatives exist in the case of health insurance as well.

Examples of self-insurance measures that reduce ill health losses are: monitoring one's health conditions to achieve early detection of serious illnesses, which ameliorates their severity; investing in medical knowledge to complement remedial medical care efforts; and making use of medical savings accounts to reduce the burden of high out-of-pocket costs. Examples of self-protection measures that reduce the likelihood that illness strikes are: employing a routine of diet, exercise, a myriad of safety measures and life-style choices, as well as using preventive medical care services (annual checkups) to monitor threats to health. Clearly, some of these measures may reduce both the probability and severity of illness.

Self-insurance is intrinsically a substitute for market insurance whereas self-protection could in principle be a substitute or a complement, depending on whether insurance companies monitor individual efforts at self-protection and reward such behaviors 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 effective, increase the likelihood of a "corner solution" in which the purchase of insurance is eschewed altogether. This possibility, which we highlight in the next sections, has been entirely missing from the debate about the rationale for mandating uniform 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 the Patient Protection and Affordable Care Act (ACA) and assess some of its welfare implications.

How relevant is this omission?

We attempt to answer this question by formulating a "full insurance" paradigm for health insurance that recognizes the alternatives of self-insurance and self-protection (SISP) and use it to address what we call *the problem of the uninsured* via two essentially nested models: a baseline model in which all losses are monetary and utility is just a function of income, or consumption (section 3), and a more comprehensive model in which utility is enhanced by both consumption and health as complementary commodities, and losses from ill health lower individual welfare as well financial income (section 4). Both models capture a special feature of the health insurance system – the role of our parallel "safety net system". This system includes voluntary medical services by charitable institutions, emergency hospital rooms, and physicians' clinics, which provide emergency services free of charge to all self-declared indigent patients due to the Hippocratic Oath, or charity. The safety net system thus amounts to a special example of "self-insurance" which allows for free riding. Our model permits an assessment of the degree to which the system contributes to "the problem of the uninsured."

We apply our calibrated "full insurance" paradigm to address four related questions:

A. What contributes to the decision not to insure and its numerical dimensions? Specifically, of the current proportion of the uninsured, what fraction would be voluntarily uninsured under 4 hypothetical scenarios: when only market insurance is available; when market insurance and both self-insurance and self-protection are available; when market insurance and just the safety net are available, and when market insurance and all the 3 other alternatives are available.

B. What is the quantitative importance of self-insurance and self-protection efforts and their impact on health losses relative to market insurance? What is the role of other key determinants of the full insurance decision in determining the optimal composition of the specific components of full insurance, including the decision "to insure or not to insure".

C. How effective would be the mandated ACA level of sanction or "tax" to be imposed on the uninsured in inducing the currently uninsured to become insured at current premium levels? In this context we attempt to estimate the "take-up", or compliance rate that would be induced by the individual mandate of the ACE based on the "full insurance" models, and compare it with alternative estimates offered by the CBO, which neglect the role of SISP (section 5).

D. What would be the net effect of the individual mandate on the overall level of medical care spending and health benefits to be derived by the previously uninsured – the main stated objective of the ACA reform plan (section 6)?

The comprehensive full insurance model and its application to health insurance inevitably involve a number of simplifying assumptions to facilitate numerical solutions. But the model is sufficiently general to allow for calibrated simulations which successfully simulate key empirical data concerning health insurance. The simulations indicate that the omission of SISP as alternatives to market health insurance may grossly overstate the success of the mandate or "tax" provision of ACA to achieve its intended compliance rate; and that at least some welfare objectives of the mandated insurance scheme may fall short of their intended goals.

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: income smoothing across different states of the world, loss reduction, and loss prevention. In the binary case having just two states of the world, 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, $L(L^e,c)$ and $p(p^e,c)$, were effective and convex such that $L'(L^e, c=0)$ and $-p'(p^e, r=0) \rightarrow \infty$, SI and SP, and hence the optimal full-insurance decision would not be subject to a corner solution. The optimal market insurance component, however, could be nil for people in a heterogeneous population with different sets of characteristics including endowed probabilities of illness, income, and attitudes toward risk. This would be the case if the market price of insurance, representing the terms of trade between income in the two states of the world, were fixed at a level $\pi^0 = [(1+\lambda)p^0/(1-p^0)]$, dictated, say, by the average odds of loss for the insurance pool $p^0/(1-p^0)$ and an insurance loading factor, λ, but did not reflect differences in individual endowments or efforts at self-protection. This assumption is invoked in the following analysis, since it reflects the structure of a typical health insurance policy which is based on community rating. The fixed price level π^0 could therefore be seen by many individuals with varying characteristics and behaviors as deviating from their own actuarially fair values.

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

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

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, $UU(p^e)$. The condition for a corner solution is that π^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 insurance loading term is high (see Fig. 1).

How would the insurance choice change if self-insurance and self-protection (SISP) were feasible alternatives?

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 in Figure 1.[3] SI would always be adopted if the absolute slope of TT, at point E, $-1/[L'(c) + 1]$, were 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, S_1. If the slope (π^0) of the market insurance budget line passing through S_1, $MM(p^0)$ were steeper than that of TT, however, it would also be steeper at point E because of the convexity of TT. Self-insurance would then completely "crowd out" market insurance. The condition for j eschewing MI is thus:

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

where the LHS of (2) represents the slope of the self-insurance transformation curve, TT.

b. *Self-protection*. Assuming that individual self-protection offers a similarly effective and convex technology for loss prevention, SP too would always be adopted, 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 π^0, 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 then shift from S_1 to S_2 – the new tangency position between TT and the highest attainable indifference curve UU(p*), generating a reduction . Market insurance remains nil since the slope of the transformation curve becomes even lower relative to π^0, i.e.,

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

where c** and (r*) denote the optimal SI and SP at point S_2. This analysis can be summarized by the following propositions:

(a). ***"****Full insurance" can never be zero if the technologies of self-insurance and self-protection are effective and convex, such at $-L'(c=0)$ and $-p'(r=0) \rightarrow \infty$.*

These conditions guarantee at least some positive spending on SI and SP.

(b). *SI and SP are joint substitutes for MI; an improvement in the technologies producing both raises the likelihood that MI is eschewed when the net price of insurance is fixed.*

Graphically, improvements in the SISP technologies make both the transformation curve and the indifference system flatter, relative to market insurance budget line. Such improvements ultimately make the absolute slopes of both curves lower than that of the market insurance budget line, π^0, at the endowment position.

(c). *For a given amount of MI, SP (c*) & SI (r*) are substitutes. If more SP* is used, SI* falls.*

This is easily seen if market insurance is nil, since an improvement in the SP technology necessarily moves the optimal self-protection leftward on the transformation curve TT.

(d). *For risk averse consumers, the last dollar optimally spent on self-protection (r*), relative to self-protection (c*), has a larger proportional impact on the probability of loss (p*) compared to the severity of los (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: see Chang and Ehrlich (CJE, 1985). The intuition is that 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).[4] This proposition offers a *corollary*:

(e). *A corollary to proposition (d) is that 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 (loss prevention), while remedial care targets largely health restoration (loss reduction).

These propositions indicate that self-insurance and self-protection may play a non-trivial role in determining health insurance choices and drawing related policy inferences. A remaining issue, however, is *how important* is this role quantitatively. We address this issue in the following sections.

3. Baseline Model

A. Simplifying assumptions:

We consider a heterogeneous population characterized by endowed probabilities of sickness, p^e, which are uniformly distributed on the open interval (0 1). All other parameters are taken to be equally distributed (income endowments, premium, production parameters, preferences). The loss from getting sick, L^e is purely monetary; we ignore any consumption aspects associated with health insurance and take insurance to be of the 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. Both assumptions are relaxed 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. Since the policy is inherently actuarially unfair, the coverage (payout) rate, however, is restricted not to exhaust the endowed loss, so $L^a \leq L^e$, where L^a sets a maximum level of coverage to take account of losses from ill-health, like sick time, which are typically not covered by insurance. The policy sets a single premium level, R, applying to the target population.[5] 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, $A(c)L^e$. The production function governing A(c) is specified as the convex function

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

with $A(0) = 1$ and $-A'(c=0) \rightarrow \infty$, so some SI is always optimal, but $A(\infty) = A_h$, to set a limit on the effectiveness of SI in conformity with the shape of TT in Figure 1.

c. *Specifying the safety net (SN) health 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) $L^0 << L^a$.

d. *Specifying the Self - Protection Production Function*: The probability of falling ill, like its associated loss, can be lowered by a proportion B(r) to $B(r)p^e$ with B(r) specified as a convex production function of self-protection spending, r, as follows:

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

with $B(0) = 1$ 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.

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

(7) $U(I) = (I^{1-\sigma} - 1)/(1- \sigma)$

In calibrating the utility function, we restrict $\sigma = 2$, 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:

$I_1^{UN} = I_1^e - c - r$ with probability of $1 - B(r)p^e$
$I_0^{UN} = I_0^e - c - r - A(c)L^e + L^0$ with probability of $B(r)p^e$

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

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

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:

$$(9)\quad A'(c^*) = -(1/L^e)\{1+[(1-B(r^*)p^e)/B(r^*)p^e][(U'(I_1^N)/U'(I_0^N)]\},$$
$$(10)\quad B'(r^*) = -\{B(r^*)p^eU'(I_0^N)+[1-B(r^*)p^e]U'(I_1^N)\} / p^e[U(I_1^N)-U(I_0^N)],$$

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

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

while staying 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 (to exclude those who are eligible for Medicare and CHIP), and live in households with income higher than 133% of the federal poverty line (to exclude those who are eligible for Medicaid). 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 parameters of the production functions for SISP, as specified in equations (4) and (6), we set $\{A_h = 0.8, \eta_1 = 0.05\}$ and $\{B_h = 0.7, \eta_2 = 0.05\}$, respectively. According to MEPS the average (expected) level of medical expenditure is $3,300. Since the maximum indemnity coverage is restricted to be $L^a = 0.8L^e$ (to limit the loss-restoring capacity of SI), we set L^a to be $6,600 and calibrate the actual value of L^e via our simulation to be $8,250. The remaining free parameter in our simulation – the level of coverage provided by safety net system – is calibrated by our simulation to match the fraction of the uninsured population - assessed in the 2009 MEPS to be 20% of the target population. The calibrated value is thus estimated to be $L^0 = 12.73\%\ L^e$.

D. Solving the model

In solving the model numerically, we aim to achieve 4 related objectives: decomposing numerically the uninsured population by the main determinants of the decision to eschew insurance; estimating numerically the magnitudes of the main control and state variables of the model; estimating the effects of the main determinants of the full-insurance decision numerically and via comparative statics; and confirming the consistency of the results with the main propositions of section 2.

a. *Decomposing the uninsured* (Table 1)

Calibrated to match the fraction of the target population that is uninsured at 20% we are able to decompose the latter into three major factors driving their decision. That is, of the 20% uninsured we estimate that 50.3% are motivated by the availability of the three alternative forms 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]. The remaining 49.7% of uninsured can be explained by highly "unfair" price of market insurance as viewed by the uninsured.[6]

b. *Quantifying the importance and relative impacts of self-insurance and self-protection*

Our calibrated simulations can be used to illustrate the quantitative importance of SI and SP outlays, c* and r*, as components of the full insurance decision under two options: being insured as opposed to uninsured. Table 2 shows the pattern of the results.

First, consistent with proposition (b) 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 a complement 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 (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 are also consistent with 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 the reductions in the endowed sizes and probabilities of loss associated with the 5 hypothetical levels of endowed risks. As indicated by the ratios of the optimized relative to the endowed magnitudes, both 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 achieved by optimal self-insurance, in conformity with proposition (d) in section 2. The same holds for the relative comparative 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 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.

Our calibrated simulations also verify that SI and SP are substitutes, in line with proposition (c) 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. *Quantifying the role of other key determinants of the full insurance decision*

The calibrated simulations of the baseline model also provide insights about the implicit role of the price of insurance in explaining the problem of the uninsured. The latter is defined in section

2 to equal $\pi = (1+\lambda)p/(1-p)$, where λ denotes the insurance loading factor which indicate the deviation of the actual insurance price relative from its actuarially fair value p/(1-p) (see Ehrlich and Becker 1972). The price π can be shown to be equal to (coverage – premium)/premium. Using MEPS data about average insurance coverage and premiums in the target population we calibrate $\pi^0 = (6600-960)/960 = 5.875$. At an endowed probability value of 10%, therefore, the loading factor becomes $\lambda = (90/10)/5.875 - 1 = 53.2\%$.

More generally, we can recover the value of the endowed probability of ill-health – the source of individual heterogeneity in the baseline model – 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 $\bar{p} = B(r^*)p^e = .143$. 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 in the Appendix.

4. Extended Model

A. Relaxing key limiting assumptions:

The baseline model enables an assessment of the role of the three components of the full insurance decision in explaining the problem of the uninsured within the more conventional framework of insurance, where smoothing income fluctuations due to unforeseen medical 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 is intended to recognize health and ordinary consumption to be complementary goods, and “health smoothing” to be the basic objective of health insurance. The latter can be achieved through insurable remedial care services that can 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. It also exposes the role that a typical health insurance policy can play in affecting one’s chosen level of health care services: by allowing health insurance to reimburse consumers for their actual health spending, as is typically the case, insurance coverage can therefore generate ex-post moral hazard or "overconsumption" of medical care, which we have abstracted from in the baseline model by assuming indemnity-type insurance. How would these extensions affect our assessment of the role of SISP in contributing to the problem of the uninsured and its policy implications?

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

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

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

The technologies governing self-insurance and self-protection – in this model lowering the probability and severity of illness strictly via self-efforts – remain the same as in the baseline model (see equations 4 and 6). Below we highlight the opportunities for health control enabled by insured remedial health-care services.

a. *Specifying the opportunities to remedy health losses via medical care*: If illness strikes, endowed health, H^e, is subject to a potential loss of L^e. The loss can be reduced via individually controlled self- insurance efforts to a level $A(c)L^e$, as in the baseline model. It can also be remedied, however, via insurable 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.[7] The production function linking restored health 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 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, but 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 here allow the uninsured to purchase additional health care services out-of-pocket, to further reduce their health loss 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) $\kappa P_m M + X + c + r + R = I_0^e$,

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

(14) $M^* = [\log(L^a)+\log(\eta_3)-\log(P_m)-\log(\kappa)-(1-\theta)\log(H^*/X^*)]/\eta_3$
(15) $X^* = I_0^e – \kappa P_m M^* - c – r – R$.

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

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

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) $M^* = [\log(L^a-L^0)+\log(\eta_3)-\log(P_m)-(1-\theta)\log(H^*/X^*)]/\eta_3$
(19) $X^* = I_0^e – P_m M^* - c – r$.

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

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

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) $EU^N(c,r)= B(r)p^eU_0^N + [1 - B(r)p^e]U_1^N$,

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) $A'(c^*) = - (1/L^e)(\{[(1\text{-}B(r^*)p^e)/B(r^*)p^e]U'_{1X}{}^N\text{-}U'_{0X}{}^N\}/U'_{0H}{}^N \text{-}L^a\varphi'/Pm)$,
(21) $B'(r^*) = - \{B(r^*)p^e[(U'_{0H}{}^N\varphi'/Pm)+U'_{0X}{}^N]+[1\text{-}B(r^*)p^e]U'_{1X}{}^N\} / p^e(U_0^N - U_1^N)$,

where the subscript s stands for the state of the world, i.e., s= {0,1}, and $U'_{sH}{}^N = \partial U_s^N(H,X)/\partial H$, e.g., denotes the partial derivative of U_s^N 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. The choice of being either insured or uninsured would be resolved by comparing the values of the last two terms. 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, $B_{h,}$, η_1, and η_2 defining the utility and production functions of SISP. In addition, we set the coinsurance rate at 25%, as is common in the literature, and the relative price of medical care 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 η_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**.** *Decomposing the uninsured* (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.[8]

b. *Quantifying the importance 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 3, 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 doubles for the insured but also rises for the uninsured.

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 is the recognition of health as a distinct and 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 (d) in section 2.[9]

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 14.3% - practically identical to its value in the baseline model. In addition, the calibrated uniform price of insurance π^0 and its varying net loading terms remain identical to those derived in the baseline model.

Under the reimbursement-type insurance allowed for in the extended model, medical care spending substantially exceeds expenditures 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 expenditures would rise by 361% and 389%, respectively, when the endowed risks of illness rise from 0.1 to 0.5.

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 will succeed in achieving one of the central objectives of the mandate - inducing the uninsured to become insured, i.e., to comply with the mandated provision of the reform act. To reinforce compliance, the ACA includes a sanction, or "tax", of $695 to be imposed on those in the target, non-elderly adult population who may choose to stay uninsured and avoid paying the required premium. To facilitate enforcement, the ACA requires all tax payers to indicate on their annual tax report whether they are enrolled in an accredited insurance plan. The IRS has been charged with monitoring and enforcing compliance. Upon discovery of non-compliance the IRS is authorized to impose the sanction.

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 motivation for avoiding the costs of insurance and their possible impact on the "take-up" decision, or compliance prospects, by the previously uninsured. Our calibrated simulations of both the baseline and extended "full-insurance" model offer some direct insights into this issue.

a. Experiment design

We incorporate into both our baseline and extended model the expected size of the sanction imposed on non-switchers and rerun our calibrated simulations to estimate its implications for the "full-insurance" decision, and thus the expected degree of compliance by the currently uninsured. The first issue we need to settle in this regard is how to assess the magnitude of the *expected* sanction. The issue is relevant since in practice at least some of the uninsured will be able to avoid paying the sanction because of 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 consider 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.[10]

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 penalties assure compliance?

b. Results (see Table 7):

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 remaining fractions of the target population remaining uninsured despite the imposed 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 in the extended model are thus lower. Yet according to both models, 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 into the ACA in response to the imposed mandate. 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 states that the post-policy uninsured nonelderly would become 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 alternatives to market insurance addressed in this paper. This suggests that estimates of the increase in the insured population, such as those provided by the CBO may indeed be significantly overstated.[11]

6. The Insurance Mandate's Effects on the Health Benefits and Health Care Spending by the Previously Uninsured

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 – improving the health status of the uninsured. It is arguable, of course, that viewing the decision to be uninsured as a rational choice cannot improve the individual welfare of the previously uninsured who would be induced by the force of the sanction to purchase health insurance policies previously eschewed in favor of alternative spending. But the question remains as to whether the switch to insurance status will improve the health benefits to the previously uninsured and possibly to society as a whole because of the social benefits from a healthier population. Viewing the decision to eschew market insurance in the context of the full-insurance decision, however, raises questions about even the net benefit to health and its associated costs. This is because of possible tradeoffs between greater spending on medical care services and alternative efforts to maintain health through self-insurance, self-protection, and out-of-pocket spending on medical care services, as well as reliance on free albeit limited services provided by the safety net for both those who decide to switch to insurance status and those who choose to remain uninsured.

Both the baseline model and the extended model can provide insights into this issue through calibrated simulations since both models solve the full-insurance decision. However, the baseline model has an important limitation in this regard because it does not model 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 expected illness losses (net of spending on self-insurance and self-protection) while those staying uninsured experience a very minor decrease in such losses.

The extended model, however, is designed to account for all of the four measures of the full-insurance decisions and the control variables of the model affecting one's health: self-insurance (c^*), self-protection (s^*) and medical care outlays M^* - both insured and out of pocket - through the production functions (4), (6) and (12). Table 8 reports the net effect of the sanction on the health and medical care outlays of the previously uninsured.

The simulations show that those becoming insured would lower their SISP spending (here assumed for convenience to involve strictly self-efforts such as diet and life-style changes) since the model shows them to be jointly substitutes for market insurance. This results in reduced preventive and remedial health benefits. Even though they may no longer rely on the safety-net, as we assume in the extended model, they now pay for it 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). 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 which the literature has identified as "ex-post moral hazard". The higher spending on medical care services, however, would lead to improved health benefits in our extended model, since the increase in all remedial care inputs is assumed to produce positive health benefits by equation (12). The net effect on health benefits enjoyed by the switchers is found out to be favorable.

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 as well as their optimal out-of-pocket spending on medical care, and increase their reliance on the safety-net system for medical care services. The net effect of the mandated health insurance on the health benefits enjoyed by members of this group, as estimated by our calibrated simulations, is unambiguously unfavorable.

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 8 reveals, the favorable net effect on the switchers outweighs the unfavorable effect on those remaining uninsured, essentially because the switch to insured status substantially increases the consumption of health care services in absolute terms or as a percentage of the expected losses from ill-health, $EL^* = (p^*)L(c^*)$.[12]

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 net effect on the expected utility of all this is unfavorable. This result follows predictably from our basic economic approach since all members of the previously uninsured group could have previously chosen to become insured but decided in favor of the no-insurance option.

7. Conclusion:

The thrust of this paper has been that a better understanding of the decision to eschew health insurance can be gained by considering it in the context of the more relevant "full insurance" problem, which includes self-insurance, self-protection, and the safety-net system (a special case of self-insurance) as alternatives. Having gone through the latter problem theoretically and via calibrated simulations, we can now return to the question we posed in the introduction: how important is the addition of the alternatives to market insurance for understanding the problem of the uninsured. Our theoretical analysis and calibrated simulations of both the baseline and extended models developed in this paper indicate that the addition is substantial.

The reason is both methodological and practical. Adding self-efforts as alternatives to the main function of insurance – improving the distribution of welfare losses across different states of the world – puts the "problem of the uninsured" in a larger context in which the specific interactions between MI, SI and SP are properly accounted for. This is especially relevant in the case of health insurance in view of accumulating scientific evidence suggesting that individual efforts to maintain proper diet and exercise, avoid hazardous consumption and related life-style choices, and spend resources on preventive health care services play a critical role in reducing the likelihood and severity of various illnesses, no less decisive than the role of the remedial care services that are rendered by the medical care system to restore health after illness strikes.

Despite the inevitably strong assumptions made in developing an operational mechanism to implement the full-insurance model empirically, the calibrated simulations of both the baseline and the extended model match quite well the empirical evidence in 2009 as provided by MEPS concerning the percentage of the uninsured and the average medical expenditures on health care services targeted by simulations. Despite their differences, both models also produce a consistent pattern of quantitative solutions for the models' key control variables and basic parameters, such as the quantitative values of self-insurance and self-protection efforts (SISP) and their interaction with medical care services, which have important policy implications as well.

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. Tables 3 and 6 indicate that optimal efforts devoted to especially SP but also to 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 expectations, 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 of 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. 12).

The analysis in section 6 also indicates that although the mandated health insurance provision may result in an overall improved health benefits for those who will decide to become insured, these benefits may fall quite short of the plan's intended outcomes, essentially because inducing the previously uninsured to become insured by the penalty imposed on the uninsured will be offset by a reduction in their previously optimal self-insurance and self-protection efforts. This is largely because of the inherently actuarially unfair structure of the mandated health insurance policies which do not allow for variations in insurance premiums across people with different endowed risks and do not compensate the insured for self-efforts that can lower the probability or severity of losses incurred from falling ill. Moreover, this "ex-ante moral hazard" problem is magnified by an additional "ex-post moral hazard" associated with the reimbursement feature of health insurance policies. The latter feature provides an incentive for the insured to over-consume remedial health services as the effective price for the insured is just the payout, or coinsurance, rate of the market price.

As Table 8 illustrates, the net increases in health benefits for the switchers are quite modest. Moreover, the added health benefits are brought about primarily through a much larger increase in the consumption of insured medical care services, which is driven by the ex-post moral hazard and would exert greater pressure on scarce medical resources.

Furthermore, our calibrated simulations in Table 8 indicate that the improved health benefits to the previously uninsured who choose to become insured is more than offset by an even larger reduction in their regular consumption benefits. This is also the case for those who choose to stay uninsured despite the imposed sanction costs since they will experience a reduction in both health and consumption benefits. The expected utility loss to all the previously uninsured consumers, whether they choose to insure or not to insure, runs contrary to the avowed objectives of improving the lot of the uninsured, but is predictable by standard economic theory. This is because the decisions to switch or stay put are induced by the penalty imposed on non-switchers, rather than by an independent choice by the uninsured which was available to them when they chose to stay uninsured.

These estimates are based on many simplifying assumptions underlying our models, as well as imperfect data we use in our calibrated simulations. 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 a denial of access to the markets for either health insurance or health care services. The policy or welfare implications of our calibrated simulations could also be made more accurate if we stratify our target population by income and age instead of just endowed probabilities of incurring health hazards. These can improve the accuracy of our estimates but may not change the thrust of our results. The central message of this paper is that useful analyses of the problem of the uninsured should consider the role of self-insurance and self-protective in deriving policy implications for alternative health reform plan.

References

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

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.

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

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.

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

Table 1: Decomposing the Uninsured: Baseline Model

Feasible "insurance" and "protection" 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 calibrated parameters for the baseline model as follows:

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

$A(c) = A_h + (1-A_h)\exp(-\eta_1 c)$ 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 note 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: Decomposing the Uninsured: 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 calibrated parameters for the extended model as follows:

$\sigma = 2$; I = \$36,000; $H^e = 11{,}300$; $L^e = 0.5\% H^e$; $L^a = (.8)\, L^e$; $L^0 = 33.6\%\, L^e$

R = \$960; $\kappa = 25\%$; Pm = 1.75; $\theta = -7$; $\eta_3 = 0.00183$

$A(c) = A_h + (1-A_h)\exp(-\eta_1 c)$ 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 and 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.

Table 8: Changes in Health and Health Spending under the Health Insurance Mandate: Extended Model

Group	**Change in expected health[1]**		**Change in expected medical care spending[3]**		Change in expected utility ($x10^{-10}$)
	Change	% of expected loss from sickness[2]	Change in $	% of expected health expenditure	
Penalty Level 1: Penalty as 72.4% of premium[4]					
Switchers	0.01936	0.48%	246.21	64.07%	-3.04
Non-Switchers	-0.00108	-0.17%	2.32	-2.83%	-5.88
Penalty Level 2: Penalty as 47.4% of premium[4]					
Switchers	0.02358	0.48%	290.69	61.53%	-1.99
Non-Switchers	-0.00114	-0.08%	-3.06	-1.84%	-3.76
Penalty Level 3: Penalty as 23.2% of premium[4]					
Switchers	0.02741	0.48%	330.55	60.27%	-1.03
Non-Switchers	-0.00077	-0.04%	-2.28	-0.94%	-1.91

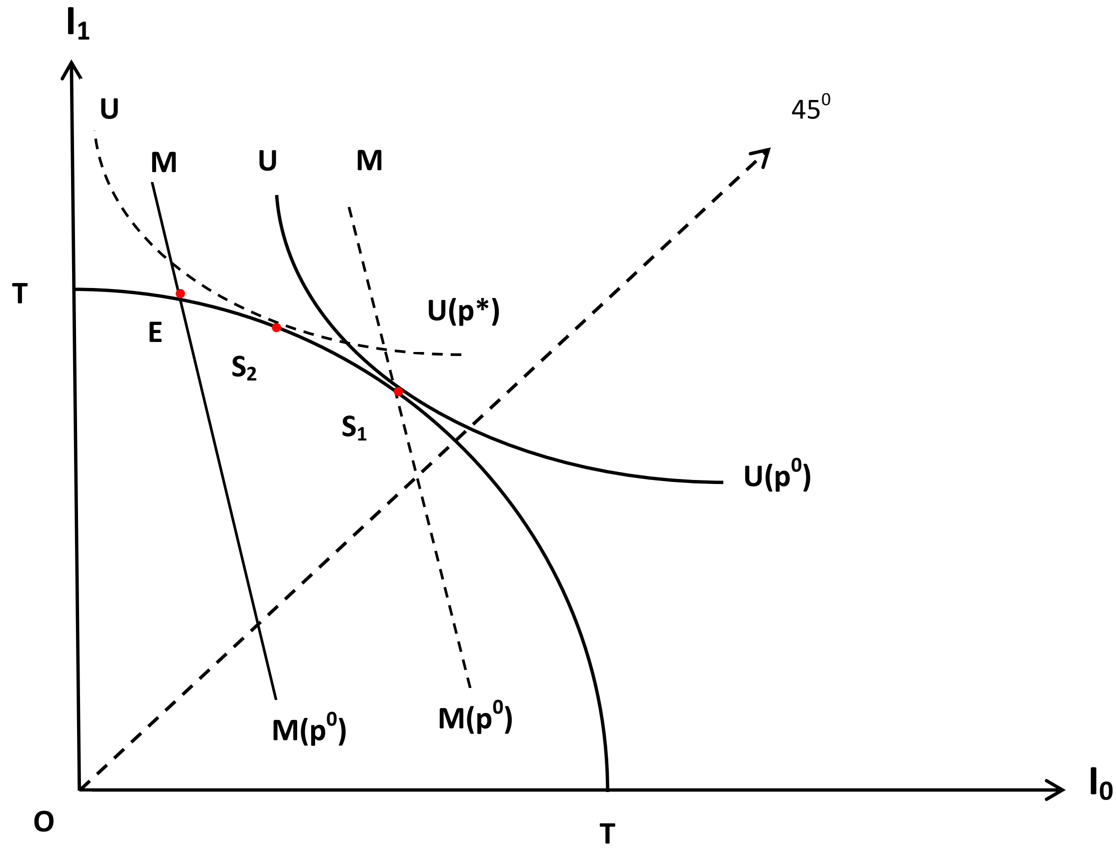
[1] Expected health is calculated as $EH = [1-B(r^*)]H^e + B(r^*)[H^e – A(c^*)L^e + L^a\varphi(M^*)]$ for the insured and $EH = [1-B(r^*)]H^e + B(r^*)[H^e – A(c^*)L^e + L^0 + (L^a-L^0)\varphi(M^*)]$ for the uninsured.

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

[3] Expected health expenditure is calculated as $B(r^*)PmM^*$.

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

Figure 1: Rationalizing No-Insurance



Appendix: Estimating the role of the major determinants of the full insurance decision via Comparative Statics

A1. The Baseline Model

Table 1 summarizes the qualitative effects of upward shifts in the main parameters of the model on the full insurance decision broken into three components: a. the impact on whether to insure or not to insure; and the impact on the full insurance choice conditional on b. choosing to insure; or c. choosing not to insure. Note that under indemnity insurance coverage and fixed premiums there is no effect on the demand for market insurance. Table 1 therefore reports just the impacts on optimal self-insurance and self-protection by the insured and uninsured groups.

Wealth effects:

As pointed out in EB (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. Perhaps in the realistic case, a higher endowed wealth increases the potential exposure to risk (monetary loss) as well, since 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 chooses to remain uninsured, and also raises optimal spending on SI (c*) and SP (r*).

Endowed loss from sickness: the results are identical to those for the "neutral" income effect.

The indemnity size: It provides a bonus to the decision to insure, but the results for optimal SISP by the insured are ambiguous: since the change amounts to a higher endowment in state 0, this 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.

Size of the premium: 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, $\partial A(c)/\partial c = -(1-A_h)\,\eta_1 \exp(-\eta_1 c) < 0$. A lower A_h would thus result in a higher marginal productivity of c at 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 elasticity of each with respect to their respective arguments. The results could be ambiguous also because of the interaction between SI and SP as substitutes.

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 components: a. the impact on whether to insure or not to insure; and the impact on the full insurance choice conditional on b. choosing to insure; or c. choosing not to insure. The relevant effects now concern optimal medical expenditure as well as optimal self-insurance and self-protection by the insured and uninsured groups.

Income effects

Higher income 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 their 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 uninsured would not be affected.

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

Impact on medical Expenditures: The impacts of upward shifts in the basic determinants of the full insurance decision on medical care services, M, generally go in the same direction as those on self-insurance, except for the cases of shifts in the level of safety-net services, L^0, and the relative risk aversion coefficient σ because of the interaction 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	-	+	+	-	+	+	-

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] See the 2010 study by LIMRA http://www.latimes.com/la-homeauto-story1,0,3900799.story

[2] See http://www.usatoday.com/money/perfi/insurance/2010-12-03-1Alifeinsurance03_ST_N.htm

[3] The shape of TT 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.

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

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

[6] 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 **(a)** in section 2.

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

[8] 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 **(a)** in section 2.

[9] The substitution relation between self-insurance and self-protection is verified in the extended model as well, in conformity with proposition (c) 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.

[10] 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.

[11] 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 taking 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 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.

[12] In this analysis we focus on the mandate's implications for the previously uninsured, and do not include any implications for the previously insured, essentially because it is not clear how they would be affected by the mandated reform plan. If the increased insurance premiums paid by the switchers is transferred directly to the previously insured, the positive income effect on members of this group will also increase their spending on self-insurance and self-protection which by our comparative statics analysis in Appendix), but this direct "income transfer" is doubtful since all taxpayers, not just the previously insured, may be the beneficiaries.