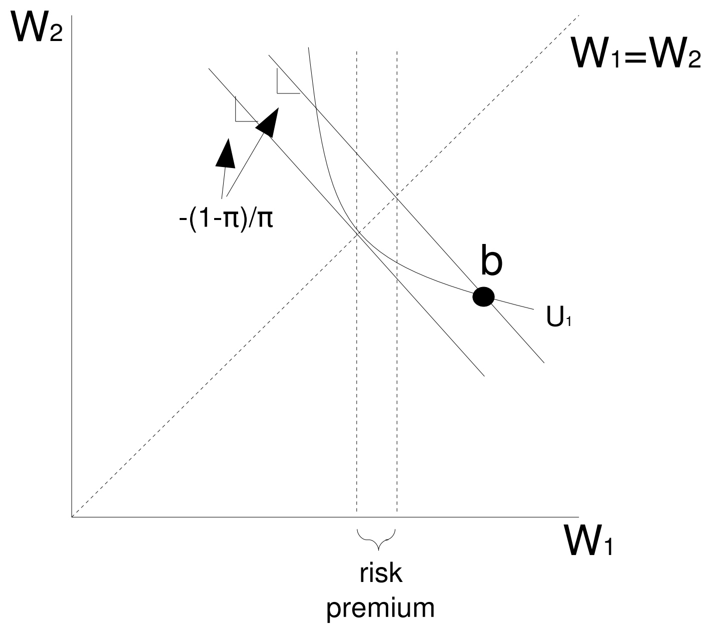
Adverse Selection and Moral Hazard in a Model With 2 States of the World

• A model of a risky situation with two discrete states of the world has the advantage that it can be neatly represented using indifference curve diagrams, with standard convex indifference curves representing risk averse consumers (see diagram below, showing the risk premium for risky prospect *b*). This enables a simple model of market failure due to adverse selection in competitive insurance markets (Rothschild and Stiglitz, 1976), and the efficiency loss due to moral hazard resulting from a situation with unobservable effort and a risk averse agent, to be constructed and illustrated graphically.



• First, we cover some conceptual preliminaries: Assume that there are 2 states of the world, a bad state (state 2) and a good state (state 1), with probability of the bad state occurring being π. Assume that individuals have a utility function $g(W_i)-e$ which gives utility in state of the world *i* given wealth in state *i* of W_i and effort *e*. Note that effort is not relevant to the adverse selection model, and that the function *g()* is the same in all states of the world, except for its input W_i, thus satisfying the axioms of expected utility theory. Individuals can only differ in the probability π they face. Suppose there are two types of individual, high risk and low risk (and note that *low risk* individuals are *high effort* individuals in the moral hazard model- their higher effort *reduces* the chance of the bad state of the world occurring). Hence the two types have the following expected utility functions, with $\pi_H > \pi_L$:

$$U_H = \pi_H(g(W_2)) + (1-\pi_H)(g(W_1)) - e_H$$
$$U_L = \pi_L(g(W_2)) + (1-\pi_L)(g(W_1)) - e_L$$

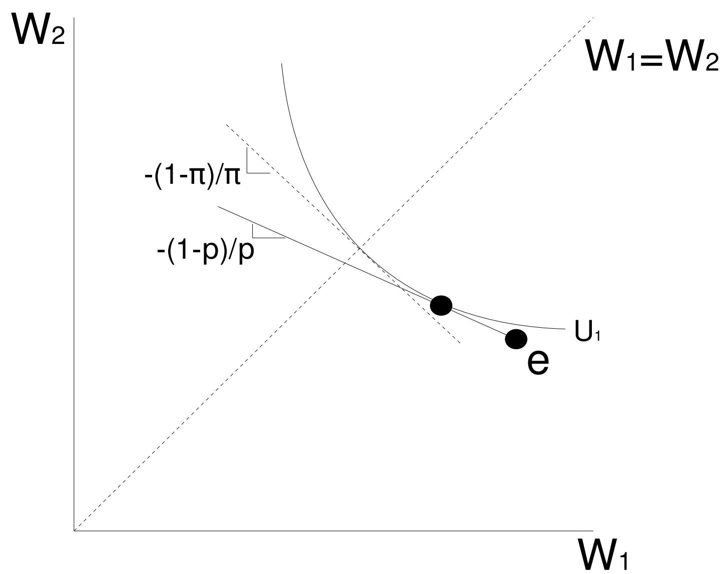
• An important result is that the marginal rate of substitution (the gradient of the indifference curves in W_1,W_2 space) of the two types is then given by:

$$MRS_{1,2}^{H} = -\frac{\left(\frac{\partial U_H}{\partial W_1}\right)}{\left(\frac{\partial U_H}{\partial W_2}\right)} = -\left(\frac{1-\pi_H}{\pi_H}\right)\frac{\left(\frac{dg}{dW_1}\right)}{\left(\frac{dg}{dW_2}\right)}$$

$$MRS_{1,2}^{L} = -\frac{\left(\frac{\partial U_L}{\partial W_1}\right)}{\left(\frac{\partial U_L}{\partial W_2}\right)} = -\left(\frac{1-\pi_L}{\pi_L}\right)\frac{\left(\frac{dg}{dW_1}\right)}{\left(\frac{dg}{dW_2}\right)}$$

Along the 45° $W_1=W_2$ line, the MRS becomes $-\frac{1-\pi_H}{\pi_H}$ and $-\frac{1-\pi_L}{\pi_L}$ respectively. This reflects the important fact that the high risk types always have shallower indifferences curve than the low risk types. (Intuitively, because the bad state of the world is more likely for them, they do not require so much additional wealth in the bad state in order to give up wealth in the good state.)

• Also, note that if either type is offered insurance at premium rate *p*, in other words offered the chance to pay *pX* in both states of the world in order to get coverage of *X* in the bad state of the world then they are effectively being offered an exchange rate between W_1 and W_2 of $-\left(\frac{1-p}{p}\right)$. Hence, if $p=\pi$ then each type will fully insure by picking an optimal bundle on the 45° line. If $p>\pi$ then an individual optimally chooses to partially insure by picking an optimal bundle to the right of the 45° line (this is illustrated below).

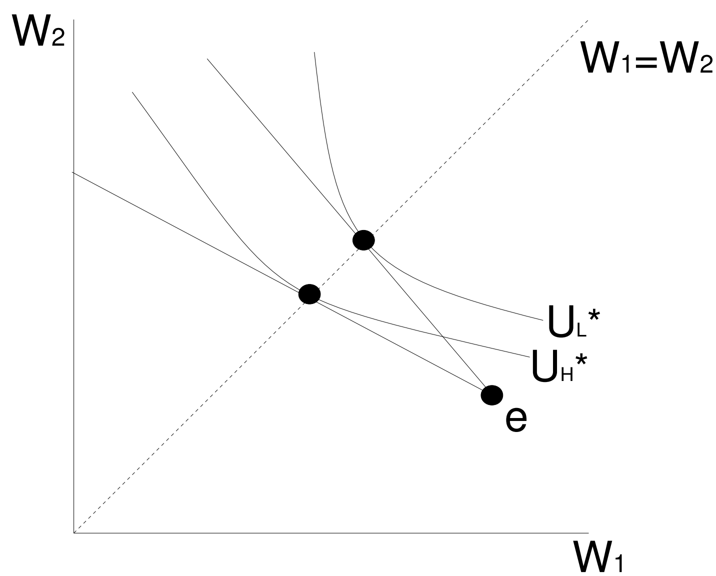


Adverse Selection Model (Rothschild and Stiglitz, 1976)

•In the adverse selection model, agent's effort has no impact on risk, and so for simplicity we can set $e_H = e_L = 0$. The two types differ in π in a manner that is beyond their control.

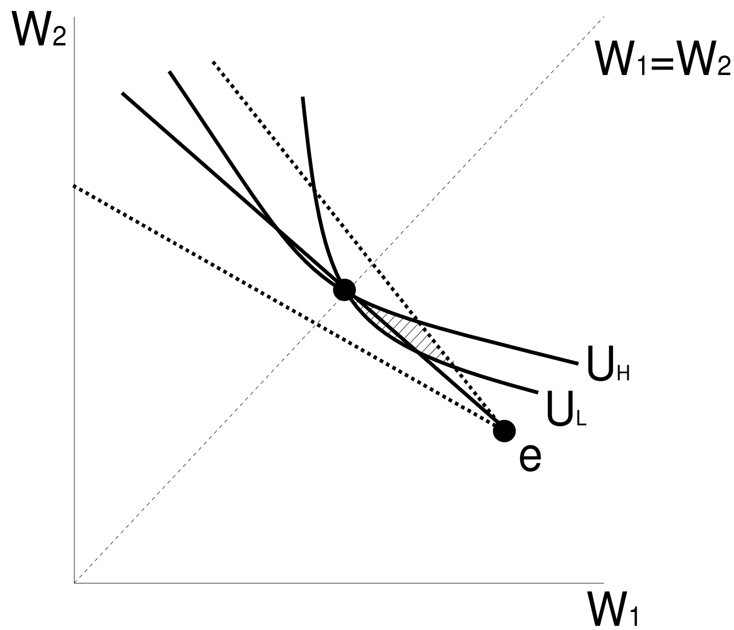
•Suppose that insurance firms play a simultaneous game in which they offer insurance contracts (each possible contract being a single point in (W_1,W_2) space), and the different types of individuals then select their preferred contract. Firms aim to maximize their profits given the response of their customers and the contract(s) offered by their competitors. Clearly, consumers pick the contract offered which lies on their highest indifference curve.

•We first consider the situation in which the insurers can distinguish between high and low risk types. In that case, each type can be offered actuarially fair insurance with their premium p tailored to their type. Perfect competition between insurance companies will force premiums to be actuarially fair in equilibrium, and will also force firms to offer full insurance. This is illustrated in the diagram below. Note that the endowment point e lies to the right of the 45º line because without insurance wealth is higher in the good state of the world than in the bad. The steeper diagonal line represents the fair insurance line for the low risk types, the shallower one the fair insurance line for the high risk types.

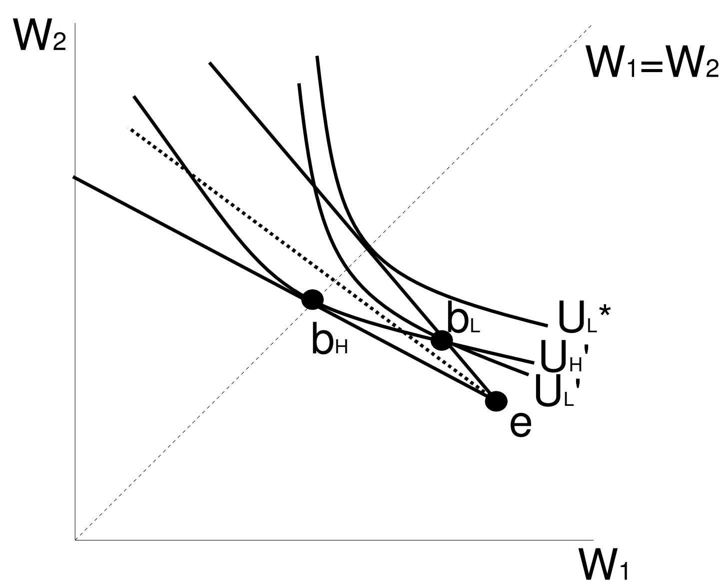


•Consider instead a situation of adverse selection where individuals know their type, but firms do not. Clearly, high risk types will now have an incentive to conceal this information from the insurer and imitate low risk types, unless the insurer can design a contract which will lead them to reveal their type.

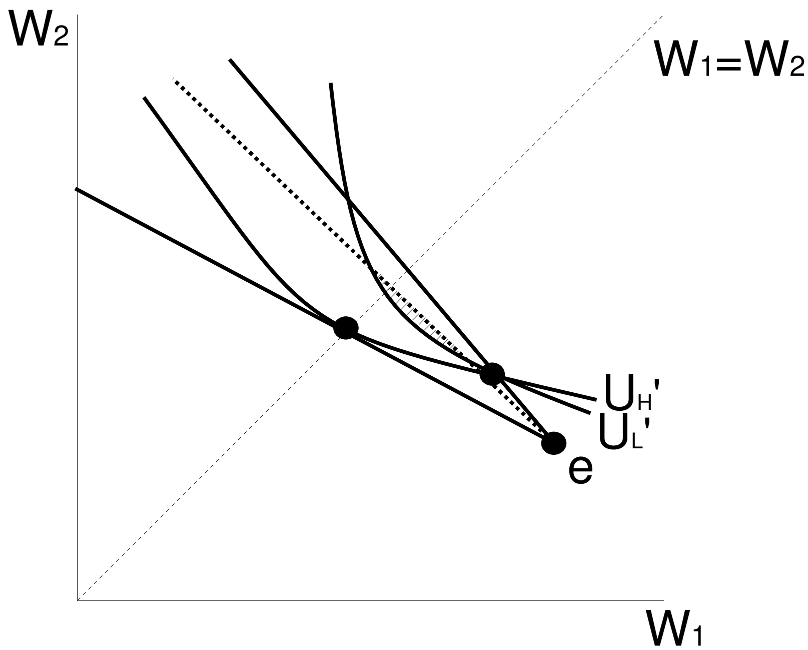
•This might, at first glance, be unnecessary. A **pooling equilibrium** would involve each firm offering a single contract based on the average risk. Thus the fair pooling insurance line lies in between that for the two types in the diagram below. Unfortunately such a pooling equilibrium is not possible because there are contracts which lie in the shaded area which a competitor could offer which are preferred by the low risk types to the pooling equilibrium contract, but not by the high risk types, and lie below the low risk fair insurance line, showing that they would create positive profits. This means intuitively that other firms would be able to “poach away” the low risk types.



•Consider now whether a **separating equilibrium** could exist, and its efficiency properties. This would involve two contracts aimed at the high and low risk types, such that they will "self-select", thus revealing their type. In order to prevent the high risk types from imitating the low risk types and taking the "wrong" contract, the low risk types would need to be offered partial insurance, at point b_L, rather than full insurance as in the case with perfect information. The low risk types will therefore be on a lower indifference curve than with perfect information, and thus this outcome would be Pareto inefficient. Although inefficient, such an equilibrium will exist if the pooled insurance line passes underneath the indifference curve of the low risk type, labelled U_L' in the diagram below.



•However, if the pooled fair insurance line lies above the indifference curve for the low risk types, as will occur if there is a large proportion of low risk types, and is illustrated below, then the separating equilibrium does not exist either, because *both* types could be "poached away" by a competitor offering a pooling contract in the shaded area. In fact no stable equilibrium then exists, and the model predicts unstable competitive insurance markets.



•**Policy consequences** – This classic model suggests that adverse selection will render competitive insurance markets at best inefficient, and at worst unstable. Signalling or screening might alleviate the problem, by enabling easier separation of high and low risk types, but probably still with a deadweight loss. Also, there may be an equity problem with the separating equilibrium (both under perfect information and under asymmetric information). For example, it may not be morally acceptable to say that in healthcare, for instance, people with pre-existing conditions should pay higher premiums, or that men should pay higher automobile insurance (due to higher risk of crashing) or that women should get less generous pensions for a given contribution record (because they are at higher risk, from the insurer's perspective, of living longer). If there are both equity and efficiency problems with competitive insurance markets, this provides a justification in economic theory for various forms of *social insurance* (e.g. UK citizens are effectively forced to purchase health insurance in a "social" pooling contract by the National Health Service).

Moral Hazard due to Unobservable Effort

•A second key application of this framework would be the deadweight loss to society due to the agency cost of inducing an agent to produce an effort level when the agent is risk averse. In a moral hazard model, there is a *single* agent, who must decide whether to become low risk by exerting high effort, or high risk by exerting low effort. However, the utility functions we already defined will describe the situation provided that $\pi_H > \pi_L$ and $e_H < e_L$.

• The model would fit a number of "stories". For instance, we might have the standard textbook case of a worker (the *agent*) being contracted to perform some project for a private company (the *principal*) in which the outcome is risky but depends partly and unobservably on the effort of the worker. On the other hand, we might have the example of moral hazard in a social insurance situation. For example, whether someone is unemployed or not depends partly but not fully on the effort they put in to finding and keeping a job. In the bad state of the world, the individual is unemployed and hence has a lower wealth level than in the good state of the world. If there was no moral hazard problem (i.e. job search/retention effort was observable) then, assuming a risk averse individual, the efficient outcome would be for the government to fully insure the individual by paying unemployment benefit equal to post-tax earnings when employed. With unobservable effort, on the other hand, the optimal second-best outcome is to make wealth when unemployed lower in order to provide an incentive for high job-search effort. Hence second-best optimal unemployment benefits will not *fully* replace post-tax earnings when employed. Similarly, the second best outcome in the case of worker being contracted is that remuneration is to some degree "performance-related", even though this imposes risk on the worker, for which the contractor must compensate them.

• In what follows, we assume that the optimal second-best outcome (and therefore the profit-maximizing outcome for the contractor/employer) would require the agent to exert high effort. (The same would then be true of the Pareto-efficient outcome, since removing the agency cost would make high effort even more socially beneficial.) We will explain using the textbook case of a private company contracting a worker to perform a task. The diagram below shows the *certain* wage that would be required in order to induce the agent to take on the task and put in high effort (/become "low risk") (*point a*) and low effort (/be "high risk") (*point b*). With observable effort, the contractor would offer point a (and require high effort in the contract), and the agent would thus have utility level U_H'. With unobservable effort, the contractor could no longer offer point a, because the agent would then choose to put in low effort, and thus the contractor would not be maximizing profits. By instead offering point c, the agent would then weakly prefer to put in high effort. However, this results in a higher *expected* wage being paid, the difference being the agency cost, which conceptually is the *risk premium* required to compensate the agent for the greater risk they now face. Since the agent is on the same indifference curve as with observable effort, but the principal is now worse off, this is clearly Pareto inefficient.