

Peltzman, Sam

Working Paper

Offsetting Behavior and Medical Breakthroughs

Working Paper, No. 177

Provided in Cooperation with:

George J. Stigler Center for the Study of the Economy and the State, The University of Chicago Booth School of Business

Suggested Citation: Peltzman, Sam (2002) : Offsetting Behavior and Medical Breakthroughs, Working Paper, No. 177, The University of Chicago, George J. Stigler Center for the Study of the Economy and the State, Chicago, IL

This Version is available at:

<https://hdl.handle.net/10419/262579>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Offsetting Behavior and Medical Breakthroughs (Revised)

Sam Peltzman

Abstract

Imagine that a medical breakthrough conquers a widespread health risk. The immediate effect is that many lives are saved. But the changed health risk also affects behavior. One effect is that the time, effort and other resources previously devoted to avoiding the conquered risk will now be freed for other activities, including those that have ancillary mortality risk. Thus some of the lives saved by the medical breakthrough will be lost to other risks.

This paper argues that such offsetting risk may be important empirically. It focuses on the history of a great medical breakthrough that nearly conquered infectious diseases – the development of antibiotics and other anti-infective drugs. This advance produced a greatly accelerated decline in US mortality over a period from the late 1930s to the mid 1950s. But most of this trend-adjusted progress evaporated in the subsequent 10 to 15 years. Similar, though somewhat more muted, reversals in mortality progress occurred in most developed countries with available data.

This paper examines several different kinds of mortality data from the US and other countries. The dominant pattern in these data is consistent with an offsetting-behavior interpretation of those aggregate trends. For example, age groups that benefited most from antibiotics show the worst subsequent deterioration in mortality. That deterioration is especially visible in mortality from accidents, suicides and homicides, where, arguably, mortality can respond promptly to behavioral change. In addition, within an age group, geographical areas (States and countries) with unusually large antibiotic benefits have relatively unfavorable subsequent mortality trends.

The main exception to these patterns occurs in poor countries, but I show that this exception vanishes when we account for the more gradual diffusion of medical progress in those countries.

Offsetting Behavior and Medical Breakthroughs (Revised)

Sam Peltzman

Medical research is sometimes viewed as a series of battles against specific maladies. Progress is often incremental but occasionally punctuated by breakthroughs that make substantial inroads on the enemy in a short period of time. Economists and others also understand that the enemy is not faceless: human behavior also affects health outcomes. Thus, holding access to medical technology constant, a better educated wealthier individual will tend to be healthier. So far, however, the interaction between the technology and the behavior has been mainly ignored.

That interaction is the subject of this paper. Specifically, I explore how medical progress, especially breakthroughs, might change the choices people make in ways that affect their health. The number of choice margins affected by a significant technical advance is too large to permit unambiguous predictions about this. Accordingly, I will focus on one specific margin – the allocation of risky activities. I will show that, even at this level of detail, theory can at best highlight opposing possibilities.

The paper is essentially an empirical exploration of one of the possibilities: that risk is reallocated in ways that blunt the health benefits of medical technology. I find that this is a possibility that deserves to be taken seriously. Much of the support for that finding comes from examining mortality data before, during and after one of the great medical breakthroughs of history: the discovery of anti-bacterial drugs. The data are crude, but they suggest a surprisingly large offsetting behavioral response to this innovation.

The next section outlines some theoretical considerations motivating a behavioral response to medical advances. It is followed by the analysis of a variety of mortality data over the 60 or so years from the beginning of the anti-bacterial drug revolution.

I. Theoretical Considerations

As far as I know, the only attempt to model a behavioral response to medical advance is in Dow, Philipson and Sala-I-Martin (DPS) (1999). It is useful here as a way of bracketing the possibilities. DPS focus on the sequential nature of health risks and the complementarities that arise because of this in risk-mitigation activities. Thus, while there are many potential causes of death, an individual can die only once. This elemental fact, according to DPS, leads the individual to invest most heavily in mitigating the most currently salient risk. For example, a young person might spend relatively little on mitigating the debilities of old age and more on exercise. If the individual survives, he or she will then shift resources to mitigating risks that become more prominent as people age.

Now suppose a medical advance reduces a mortality risk that is important to the young. They still face many kinds of mortality risk in the future. According to DPS, individuals will then shift resources no longer needed to mitigate the risk that has been reduced by technology toward coping with the myriad of remaining risks.

In fact DPS – as will I – couch everything in terms of individual choice between consumption and protective expenditures. And their comparative static exercise involves a price subsidy to the latter rather than a technological advance. But I follow them in adopting a broad view of the resources being allocated. Their model has a clear

implication: if one kind of mortality risk (to other than the oldest) is reduced, there will be a favorable effect on other kinds of mortality of risk.

I get the opposite implication by emphasizing a link between mortality risk and consumption choices (and by suppressing the sequential nature of health risk.) That is, mortality can be affected by the size and composition of the consumption bundle as well as by health spending. So all three are in principle chosen simultaneously. To illustrate how a link between mortality and consumption can act to blunt the impact of medical advance on mortality, I begin with the simplest case I can think of: the individual wants to maximize expected utility (E) over a single period. To obtain utility the individual has to survive the whole period. So I ignore any utility obtained or agony endured up to the date of death. Thus the maximand is just

$$(1) \quad E = \text{probability of surviving} \times \text{Utility, given survival.}$$

As in DPS, the probability of surviving is increased by devoting resources (S) to health and safety. But the resources (X) devoted to utility-producing consumption reduce the survival probability. Total resources (S+X) are fixed.

The composites, S and X, include specific goods and services. So X would include something like driving a car, which produces both utility and a safety risk. And S would include medical expenses, which are valued mainly because they prevent illness. But they include more: time and effort can be allocated to activities that entail more risk, or to mitigating risk; life-styles can be chosen to emphasize more or less risk; government policy can emphasize medical research or building roads; etc.

One way to grasp the breadth of the relevant tradeoff is to think of what would be in X if survival were certain. There would, of course, be more of it, since there would be

no need for medical expenses. But the composition and manner of consumption of X would also change. For example, food would be richer, exercise would be less common and traffic restraints would be unnecessary. We forgo some of these pleasures because survival is in fact uncertain. Accordingly, S should be thought of as including a monetary equivalent of the utility forgone to enhance survival.

With total resources fixed, the problem boils down to picking X to maximize

$$(2) E = F(X, S) \times U(x),$$

where,

F=survival probability,

U= utility function, and

$$F_X < 0, F_S > 0, U_X > 0.$$

The first-order condition for this problem can be written:

$$(3) [(F_X - F_S) / F] + U_X / U = 0,$$

where the percentage gain in utility from an extra X just balances the percentage loss in survival probability that arises both from consuming the extra X and from having less to spend on S.

Now, suppose there is a medical advance that, for simplicity, increases the level of F. The sign of the effect of this change on the equilibrium X is the opposite of the sign of the bracketed term in (3); that is X increases. The intuition here is straightforward: the advance has reduced the percentage survival loss from the marginal X and thereby induces a balancing reduction of the marginal utility of X. The increase in X needed to

achieve the new equilibrium will engender some reduction in survival probability that offsets part of the effect of the medical advance.

The only point of this exercise is to highlight one aspect of the response to medical advances – the inducement to (risky) utility-seeking that occurs because the advance substitutes for some of the individual's own risk-reducing expenditure. Accordingly, much is being left out here. For example, by analyzing a parametric increase in survival probability, I implicitly ignore the kind of complementarities emphasized by DPS. In terms of my set up, those would translate into medical advances that raised the marginal product of S. And such advances would, all else the same, induce substitution toward more, not less, private risk mitigation expenses.

Also, the model is too sketchy and general to use as a guide to magnitudes. For the specific case analyzed above, the model does imply that the behavioral response offsets no more than half the parametric change in survival probability. But that quantitative precision is easily lost when the model or the presumed nature of the medical advance is embellished.

Finally, since there is only one source of risk, the model cannot easily handle a medical breakthrough in a specific type of risk. To do that, I expand the model to include two sources of risk. Thus, suppose there are two activities that involve consuming X and Y respectively. As before, the activities produce utility but also entail survival risk. For relevance to subsequent empirical work, think of the X-activity as producing exposure to infectious diseases (before antibiotics were discovered), while the Y-activity engenders all other risks. For example, X could involve much social interaction, while driving a car in solitude would illustrate Y. Realistically, any consumption bundle would produce

some of both risks, but for simplicity, I treat X and Y as mutually exclusive, separable activities.

As before, safety expenditures can mitigate the two kinds of risks. But here I assume that each dollar of safety expenditures affect both kinds of survival risk. Some safety expenditures undoubtedly have this general effect. For example, acquiring good habits and knowledge about health can reduce the risk of both infectious diseases and heart attacks. However, the reason I do not assume two mutually exclusive kinds of safety expenditures is pedagogical. It is simpler and permits me to emphasize a specific tradeoff.

The two-risk version of equation (2) is

$$(4) E = F(X, S)G(Y, S)U(X, Y),$$

where F is now the probability of surviving infectious diseases, G is the probability of surviving other kinds of risk, and FG is the total probability of surviving and thereby obtaining utility. So the choice problem is now to pick X and Y and, implicitly, S to maximize expected utility.

The first-order conditions for this problem can be expressed as:

$$(5) \frac{F_S - F_X}{F} = \frac{U_X}{U} - \frac{G_S}{G}$$

and

$$(6) \frac{F_S}{F} = \frac{U_Y}{U} + \frac{G_Y - G_S}{G}.$$

As before, these come from balancing the favorable effects from spending more on S – the direct effect on survival and the indirect effect from spending less on X and Y – on

the loss in utility. The favorable marginal effects of S on surviving infectious diseases have been collected on the left-hand sides of (5) and (6).

The reason for doing this is to help us think about the effects of a medical breakthrough in treating infectious diseases. Specifically, imagine an innovation that very nearly eliminates that risk. It pushes F close to 1, and thereby causes the marginal consumption hazards and payoffs to precautions in $F(\cdot)$ to vanish. This is approximately what happened after antibiotics were discovered. Mathematically, the positive expressions on the left-hand sides of (5) and (6) would both go to zero. The kind of adjustment required to restore equilibrium would entail diverting resources away from S toward X and Y – to reduce the marginal consumption utilities and raise the remaining health risk.

It is this last adjustment that I will focus on. It is not mathematically required that Y increase or increase enough to matter empirically. And allowing for different types of safety expenditures would only heighten that ambiguity. But one of the incentives created by the breakthrough is to divert resources from precaution to utility enhancing risk taking. The next section shows why we may need to take that incentive seriously when thinking about the effects of medical breakthroughs.

II. The Anti-bacterial Drug Revolution and Mortality

Infectious diseases have been a major source of mortality for most of human history. At least in the developed world, these scourges have now receded into the background. Today only pneumonia and influenza remains as one of the 10 leading causes of death in the United States, and it afflicts mainly the elderly. The contribution of

medical technology to control of infectious diseases goes back at least 200 years, to Jenner's discovery of smallpox vaccine. But there was a major acceleration of progress in the two decades from the mid 1930s to the mid 1950s.

That period began with the commercial introduction of sulfa drugs, the first broadly effective anti-bacterial drugs. This was followed by the introduction of penicillin and other antibiotics in the 1940s and 1950s as well as the development of effective anti-influenza vaccines in the mid 1940s. These had the effect of dramatically reducing mortality, especially premature mortality, from formerly dreaded diseases like tuberculosis, scarlet fever, syphilis, etc. These drugs also contributed to advances in surgery by controlling secondary infections.

Figure 1 summarizes this history. It shows the United States age-adjusted death rate¹ over the first half or so of the twentieth century. Prior to the advent of antibiotics² there had been a remarkably steady decline in mortality interrupted only by the influenza pandemic around 1918. The effect of antibiotics is evident in the lower panel of Figure 1. This shows actual mortality for 1940-1955 and a linear projection of the pre-1940 trend. There is a palpable acceleration of progress in this period: the rate of decline in mortality doubles – from around 1 per cent annually to around 2 per cent. By 1955 actual mortality is nearly 20 per cent below the extrapolated pre-1940 trend.

At this point, however, progress essentially stops. The next 15 years – from 1955 to 1970- are shown in the top panel of Figure 2. Here actual mortality is essentially trendless until the very end of the period. By the late 1960s it is almost as if the antibiotic revolution hadn't occurred, in that actual and extrapolated- pre-antibiotic era mortality nearly coincide. No previous 15-year period in the century (excluding the years of the

influenza epidemic) saw as little progress as this one. That sudden dramatic swing from above average to below-average progress is a clear motivation to this paper.

That aberrant period ends around 1970. As shown in the bottom panel of Figure 2, progress subsequently resumes. It does so at a rate comparable to the early post-antibiotic era for the next decade or so. By 1980 most of the mid 1950s gap between actual mortality and the extrapolated pre-antibiotic-era mortality is restored.

Figure 3 summarizes the twentieth century history of the age adjusted death rate with the pre antibiotic trend removed. The substantial magnitude of the antibiotic shock and subsequent rebound are clear, as is the resumption of progress after 1970. The last 20 or so years look remarkably like the first 40.

III. U. S. National Data

These aggregate trends may suggest an important role for risk-substitution. But to say more it is necessary to look behind the aggregates. In this section I disaggregate the US data by age group. Subsequent sections examine first a cross-section of states and then compare US experience to that of other countries

A. Cross-Age Group Correlations

The antibiotic revolution had uneven effects across demographic groups. The most important difference was a greater relative impact on younger age groups than the elderly. In 1938 infectious diseases accounted for around 10 per cent of deaths of those over 45. But for those younger, 25 to 65 per cent (depending on the specific age group) of mortality came from these diseases. The leading cause of death for those aged 20 to 40

was tuberculosis, which accounted for over one sixth of all deaths in this age group. Antibiotics essentially wiped out that disease.

These differences in the impact of antibiotics imply different risk-responses, if risk substitution is important. For example, we might doubt the importance of risk-substitution if most of the post 1955 retrogression is due to worsening mortality among the elderly, because they were the least benefited by antibiotics. However, the experience across age groups – summarized in Table 1 and Figure 4 – tends to dispel such doubts.

Panel A. of the table describes the experience across age groups within various demographic categories in the periods that Figures 1 and 2 suggest are interesting – the antibiotic ‘shock’ to mortality from the late 1930s to the mid 1950s, the ‘response’ (?) of little progress to the late 1960s and a resumption of progress thereafter. The data here are unweighted averages across the age groups (<1, 1-4, then each 10 years beginning 5-14). There are 9 age groups in each sample.³

The main message of panel A. is that the trends evident in aggregate mortality are broadly similar across demographic groups and characterize most age groups. There is some variety – the shock is greater for females and non-whites, for example. But every group’s mortality improves, then falls back and finally, around 1970, resumes to some degree its pre-antibiotic progress. However, the variety in panel A is not suggestive of risk-substitution – e.g., there is no difference between male and female average responses that correspond to the differences in average shocks.

The more intriguing data are in panel B, which shows correlations between the antibiotic shock and subsequent behavior across the 9 age groups. The first three columns focus on behavior from 1955-68, or just after the shock had ended. The correlations are

almost all negative, often significantly so in spite of the small sample. That is, age groups with the most favorable effects from antibiotics tended to have the most unfavorable experience subsequently.

Moreover, the relevant magnitudes are substantial. For example, the first entry in panel B (-.78) implies a response-shock elasticity only a little below -1 . This would suggest that most of any unusually favorable antibiotic shock is offset in the ensuing period. Figure 3 depicts this particular correlation, and it reveals that only 15-24 year olds depart substantially from the fitted relationship.

The other important regularity in panel B is the lower correlations in column (2) than in (1). For example, for males the two correlation differ by a factor of 2 (-.73 v. -.37). This says that the male response across ages is more highly correlated with the antibiotic shock for the whole population than with the antibiotic shock for males alone. The (correct) inference is that there is a negative partial correlation between the female shock and the male response. While the small samples warrant caution, the pattern seems too persistent to ignore.

That pattern either argues against a risk-substitution interpretation or embellishes it. If we think of gender/race groups as immutable states-of-nature, then males, for example, should not respond to the female shock. An alternative is that the female shock conveys information about consequences of a variety of life-style choices that are also available to men (and vice versa). For example, in the pre-1955 period 'female' connotes less 'blue-collar' type work than does 'male.' This alternative implies that a sub-group's response would be sensitive to other sub-group shocks, which is generally what we find in the first two columns of panel B, Table 1.

A similar argument applies to age groups. Responses to a medical breakthrough can entail life-style changes that do not affect mortality immediately. Column (3), panel B, table 1 tries to grapple with this possibility by averaging the antibiotic shock over three age groups. Specifically, this column uses an age group's 'average shock', which is just the average of the shock for that cohort and the two preceding.⁴ However, the correlations in column (3) are not distinguishable from those in (1). The data are just too crude to pursue the precise link between shock and response. For example, an appropriate average shock might include older as well as younger cohorts, because the expectation of lowered risk tomorrow should affect current behavior. However, implementing this idea also leads to results indistinguishable from the preceding.

Accordingly, the negative shock-response correlations reported above – and subsequently – should be viewed as reflections of a broad pattern. The anti-biotic shock tended to be largest for the young, smallest for the old and somewhere in between for the middle-aged. And the post-1955 mortality rebound tended to follow the same age pattern.

That broad message is perhaps better captured by Figures 5 and 6. The dotted lines show the actual cumulative mortality risk for two individuals (call them Jay and Kay) who were a year old on different dates. The smooth line is the cumulative mortality risk for a counterfactual one year old born the same year as Jay or Kay but into a world where mortality declines at the rates observed prior to the antibiotic revolution.

Jay is a year old in 1935, just in time to experience all of the benefit of the antibiotic revolution. This revolution dominates the mortality risk of Jay's youth and early middle age. So the odds that Jay will survive to any year up to 1980 are better than the counterfactual.

Kay is born in 1954, at the end of the antibiotic revolution. Like anyone born later in the century, she is destined to have lower mortality risk than Jay. Only 5 in 100 of Kay's contemporaries do not make it to age 45, compared to 7 in 100 of Jay's. However, the counterfactual Kay would have even better odds of making it to the new millennium. For that hypothetical individual, mortality risks from 1955 onward would have continued to decline at pre-revolution rates. Accordingly Kay's counterfactual twin would have obtained all the mortality benefits of the antibiotic revolution without any of the subsequent rebound. In that hypothetical world, only around 3 of 100 would not have seen the year 2000. It is that gap – more generally the uniformly positive gap between the dotted and solid line in figure 6 – that captures the response to the antibiotic revolution.

The two figures are also useful in providing a sense of magnitudes. The estimated effect of the response to the antibiotic revolution cumulates to 2 extra deaths per 100 by age 45. That is a large fraction of the actual risk of 5 in 100. But both numbers are tiny absolutely compared to what they will become when Kay's cohort become senior citizens. Even before antibiotics steady medical progress had confined mortality mainly to the elderly.

Finally, column (4) of panel B, table 1 shows that the response patterns persist into the post-1970 period, when aggregate mortality resumes its progress. Thus, the resumption of progress is more like (a series of) general advances that have roughly similar (or random) effects across age groups. The altered age structure of mortality progress evident in figure 3 persists to this day.

B. Are the Cross-Age Group Correlations Biased?

Any bias in the correlations in Table 1 would appear to be upward. To obtain the shock and response variables, I deduct a common counterfactual trend from two different trends. If everything but this common counterfactual was noise, the correlation between shock and response would be positive, not negative.

There is, however, a potential negative bias arising from the history of mortality over the relevant period. The counterfactual I have used is the trend in total mortality up to around 1940. Importantly, in that period, progress in infectious disease mortality already exceeded that in other areas. Thus the advent of antibiotics did not initiate progress against infectious diseases. But it did very nearly eliminate them as a significant source of mortality. This means that groups where infectious disease mortality is more important will tend to have absolutely larger counterfactual trends, shocks and responses, and this could induce a negative shock-response correlation that has nothing to do with behavior. The correlation would be rooted in the historical difference in the importance of infectious diseases plus the convergence of mortality rates after infectious diseases are conquered.⁵

This possibility is addressed in Table 2. It repeats parts of Table 1, but it eliminates infectious disease mortality trends from the post-shock counterfactual. That is, to estimate a group's response to the antibiotic shock, the counterfactual I use is the pre-antibiotic trend of non-infectious-disease mortality. Panel A of Table 2 shows that this counterfactual is, as expected, smaller algebraically than the Table 1 counterpart, by around .5 per cent per year on average. (Put differently, post-1955 mortality looks .5 per cent less unfavorable in Table 2 than Table 1). However, crucially, panel B shows that

the negative shock-response correlations are essentially unaffected by the changed counterfactual. Therefore something more than the potential bias from the pre-antibiotic progress against infectious diseases is driving those correlations. And we are absolved from having to choose between the alternative counterfactuals.⁶

C. What Kind of Behavior was Affected?

It is hazardous to attempt to delineate the response to a breakthrough like the discovery of antibiotics. While the model is couched in the familiar language of individual choice under full information, this is only a convenience. The affected decision margins could include such things as the politics of public funds for medical research, policies toward health education and public safety, etc. Even the individual choice margins are too numerous to catalog.

Moreover, nothing in the theory depends crucially on the chooser's knowledge of the underlying risks.⁷ What is crucial is that different utility-producing activities have different mortality risks.

These caveats understood, I attempt to be more specific in Table 3, which is based on a bifurcation of total mortality into external and non-external causes. External causes are accidents, suicides and homicides. Here, I am guessing that external causes are more immediately malleable than other types of mortality. For example, they do not require cumulative effects of changes in lifestyle.

The data in Table 3 seem consistent with that guess. The bifurcation adds some noise, and the small sample assures that none of the differences are significant statistically. Nevertheless, the immediate post 1955 response of external cause mortality

seems more negatively correlated with the antibiotic shock than the non-external cause response. The response-shock elasticities implied by these correlations differ by factors of two or three. As shown in Figure 5 (ex4), the post-1955 behavior of external cause mortality looks very much like that of total mortality in Figure 7 (ex4).

Interestingly, the direction and magnitude of these differences reverse in the post-1968 period. So the data seem to be saying that external causes responded more reliably to the antibiotic shock first, then the non-external cause response ‘caught up.’ All of the relevant correlations, however, remain negative.

IV. State Cross Section

There is enough geographic variety in the impact of antibiotics to provide further evidence on the subsequent response. For this purpose I use data from the 21 states (including DC) that had death registration systems in 1910. As detailed in the note to Table 4, this sample includes both large urban and small rural states, but none in the South. For the age groups I will subsequently use, the range of the 1940-55 annual rate of change in mortality in this sample is roughly the same magnitude as the mean change. This mean and range runs around -3 to -5 percent per year for groups under 45 and half that for the older groups.⁸

Instead of replicating the previous analysis 21 times, I estimate shock-response elasticities by age group across the 21 states. The kind of question these estimates address is: did young people in states with large (small) antibiotic shocks also experience a large (small) subsequent mortality rebound? By and large Table 4 gives a positive answer to this.

The table also addresses differences across the age groups: we have seen that the relative effect of antibiotics was greatest for younger age groups. This suggests that the strongest ‘signal’ would also be found in the cross-state variation of the shock estimates for the young. If the average shock is weak for the old, the cross-state variation would be mainly ‘noise,’ so the estimated elasticities for the old should be small. Table 4 is, however, unkind to this conjecture

The first column uses shocks and responses for a few age groups based on each state’s pre antibiotic mortality trend for that age group. The elasticity estimates are predominately negative and thereby tend to corroborate the previous across-age-group findings. However, the estimates are weak statistically.

Part of the weakness may be due to the counterfactual. It assumes that pre-antibiotic mortality improvements would be permanently confined by state borders. The more reasonable expectation is for some geographic convergence in mortality as medical knowledge diffuses, and our data are consistent with that expectation: the states with the best mortality trends tended to have the highest mortality early in the period.⁹

Columns (2) and (3) of Table 4 use a counterfactual that allows each state’s trend to converge toward a common national trend. In column (2) the rate of convergence is estimated from the data as follows: First for each age-group sample I regress the log of 1940 state mortality rates ($mr1$) on the 1910 counterparts ($mr0$) to obtain:

$$(7) \quad mr1 = a + b \times mr0 + error .$$

The parameter, b , is an estimate of the degree of convergence over the 30 year period.

It should (and does) range from zero (complete convergence) to +1 (no convergence).

Then I assume (7) holds in first differences, so the change in log mortality over any period ($dmr1$) would be related to the previous trend ($dmr0$) by:

$$(8) \quad dmr1 = k + b' \times dmr0 + error'$$

where b' is just the parameter, b , from (7) adjusted appropriately for the length of the period. The intercept, k , would be $(1-b')$ times the national trend toward which the states are presumed to be converging. The counterfactual trend for any state is then estimated by using (8) to project $dmr1$ forward from the 1940 initial conditions (the pre-1940 national trend for the age group for the intercept term and the state specific pre-1940 trend for $dmr0$.)

Column (3) imposes the same b (.5 per 15 years) on every age group, and this crude approach¹⁰ gives essentially the same results as column (2). Once I allow for the geographic diffusion of medical progress, the relevant elasticities become considerably more substantial statistically and economically. However, they are not, as I conjectured, most substantial for the younger age groups.

V. International Comparison: Was the US Unique?

From the various sources described in Table 5 I was able to piece together a sample of 19 countries with the kind of historical mortality data needed to replicate some of the preceding analysis. I divided the sample according to Maddison's (1995) estimate of per capita income in 1938, near the start of the antibiotic era and of WWII. The seven countries below .6 of US per capita income (Chile, Japan, Italy, Spain, Portugal, Ireland, Finland) are classified as 'poor' and the other twelve (Australia, New Zealand, England,

France, Germany, Belgium, Denmark, Norway, Sweden, Netherlands, Switzerland , US) are ‘rich.’

Table 5 shows summary data on mortality trends (averaged across age groups) over the relevant time periods. These make clear that the US history is hardly exceptional among rich countries. The US pattern of gradual progress that accelerates around 1940 and then is followed by a mortality rebound is the central tendency of the rich country sample. Even the relevant magnitudes are similar. The only exception is that the mortality rebound is somewhat smaller outside the US.

The poor countries also experienced gradual progress that accelerated around 1940, but there are notable differences between them and the rich countries: the poor countries’ pre antibiotic decline is smaller, the acceleration is larger, and there is no mortality rebound. Instead there is continued progress at above pre-antibiotic rates to 1970 (indeed –not shown here – to the end of the 20th century).

Part – but only part- of the rich/poor country difference is ‘catch up.’ Going forward from 1940 or so the poor countries had the benefit of the rich countries’ prior experience in what worked to reduce mortality. In addition several of them were to become rich and thereby acquire the resources to better implement that knowledge. Arguably, the pre antibiotic experience of the rich countries would be a more plausible counterfactual for the poor countries’ post-1940 experience than their own history. Accordingly Table 5 also shows summary data for the poor countries using a counterfactual based on the rich countries’ pre antibiotic era trends. That reduces, but does not eliminate, the salient differences between the rich and poor countries.

Table 6 shows the correlation across age-groups of the antibiotic shock and post-antibiotic response within each of the 19 countries. The countries are divided according to the quality of the data source. (See note to Table 6.) The top panel uses definitions and dating conventions similar to those used previously for the US (Table 2 or Figure 4)¹¹. That is, the ‘shock’ and ‘response’ are net-of-counterfactual annual changes in mortality; the shock (response) period is c.1940 to c.1955 (c.1955 to 1970); and the counterfactual is the country-specific annual change in mortality from some time before WWI to c.1940. The lower panel replaces the country-specific counterfactuals with the previously described 5 country average pre-antibiotic mortality trend.¹²

A negative shock-response correlation is typical for rich countries. This holds regardless of data quality or the counterfactual used. The correlations are somewhat sharper with the arguably more appropriate counterfactual and higher quality data. But the overwhelming conclusion from the data is that the US experience is hardly an outlier among rich countries.

The poor countries behave much differently, however. Their shock-response correlations are virtually all positive and often significant. That pattern becomes more muted with the more plausible counterfactual, but it does not go away. These positive correlations are consistent with the apparently permanent acceleration of mortality progress for the poor countries that took place around 1940. The message here is that the age-specific improvements also are long lasting. So a group that fared relatively well early on continues to do so throughout the post-1940 period.¹³

One implication of that message is that our dating conventions are inappropriate for the poor countries. Their catching up with the knowledge and income levels of the

rich countries produces a longer effective shock period. Table 7 pursues that implication. Here the shock is defined over the period ending 1970, instead of 1955 and the response period is everything after 1970.¹⁴ That re-dating produces dramatically different results. Now the poor countries begin to show very much the same negative shock-response correlation as their rich counterparts. Thus, the poor countries' response to medical advance is not, after all, much different from the rich countries'. It is just drawn out over a longer period.

Finally, Table 8 looks at the shock-response correlation within age-groups across the rich countries. The data are crudely purged of country-effects by expressing each shock or response as a deviation from the country mean.¹⁵ So these correlations ask about the response to a shock that was unusually large for the country and for the age group. For the age groups under 50 or 55 the answer is, again, that unusually favorable mortality shocks tend to be followed by unfavorable responses. Only 2 of the 23 correlations above the double line in the Table are even weakly positive. This is the same pattern we found across the US States. Unlike that sample, however, we do find here the expected weakening of the correlation at the older ages (where the measured shocks are arguably relatively noisy.)

VI. Conclusions

Many years ago (Peltzman, 1975) I argued that automobile accident mortality was affected by offsetting driver behavior following the mandatory installation of safety devices in cars. The data in this paper suggest that the perspective of that paper may have been excessively narrow. In addition to the 'within-mortality' kind of response emphasized in the 1975 paper, there appears to be a cross-mortality response that ought

to be taken seriously. This is the tendency for a favorable shock in one type of mortality to induce offsetting increases in other kinds of mortality risk.

Economic theory cannot unambiguously imply such a cross-mortality response. It can only leave it as a logically coherent possibility. The reason for giving weight to the possibility is empirical. The data I have used are crude, unfashionably aggregate and sparse. Nevertheless the important patterns in these data are all consistent with a sizeable offsetting response to the specific mortality shock studied here, the advent of anti-infective drugs.

That shock remains as the greatest achievement of modern medicine. It laid the foundation for a research-intensive pharmaceutical industry and the ongoing pursuit of similar medical breakthroughs. A young child today has essentially no risk of dying from something like scarlet fever or of growing up only to die of tuberculosis before reaching the prime of life. But, the data suggest, the absence of these long-forgotten risks has also loosened constraints on behavior that carries other kinds of mortality risk

The existence of such offsetting behavior does not diminish the importance of this breakthrough. It may however shed new light on the nature of medical progress. The more durable reductions of mortality risk may come from the steady accumulation of small advances than from the occasional great triumph.

References

- Demographic Yearbook*. United Nations, New York
- Dow, William; Philipson, Tomas and Sala-i-Martin, Xavier. "Longevity Complementarities Under Competing Risks." *American Economic Review*, December 1999, 89(5), pp.1358-71.
- Keyfitz, N. and Flieger, W. *World Population Growth and Aging: Demographic Trends in the Late Twentieth Century*. Chicago:University of Chicago Press, 1990.
- Maddison, Angus. *Monitoring the World Economy, 1820-1992*. Paris: Organization for Economic Cooperation and Development, 1995.
- Pascua, M. "Evolution of Mortality in Europe During the Twentieth Century." *Epidemiological and Vital Statistics Report*, v.3. World Health Organization, 1950
- Peltzman, Sam. "The Effects of Automobile Safety Regulation." *Journal of Political Economy*, August 1975, 83(4), pp.677-725.
- Preston, S.; Keyfitz N. and Schoen R. *Causes of Death: Life Tables for National Populations*. New York: Seminar Press, 1972.
- World Health Statistics Annual*. World Health Organization, Geneva.

Endnotes

¹¹ This is the total death rate with age-weights fixed at 1940 values.

² I will use ‘antibiotics’ as shorthand for the acceleration of technological progress that was characteristic of the period in which antibiotics were introduced.

³ I exclude the 75 and older groups where behavioral effects on mortality are arguably small.

⁴ The exceptions are the 1-4 group, which is averaged only with the under 1 group, and the latter, which is left as is. Averaging over three groups is arbitrary, but reflects the thirty-year span of the data – someone who is, say, at the top of the 35-44 group in 1970 would have been at the bottom of the 15-24 group at the start of the shock period.

⁵ The possibility can be illustrated by a highly stylized example that captures the relevant history. Suppose we have two groups whose mortality we track over three equal periods: period 1 precedes a breakthrough, and is used to establish a counterfactual trend for all subsequent periods. In period 2 the breakthrough eliminates infectious disease mortality; this produces a shock that we estimate as the group’s period II total mortality trend less the period I counterfactual. Period 3 begins when infectious diseases are conquered, and we estimate the response as the period 3 trend minus the counterfactual.

The two groups differ only in the relative importance of infectious disease mortality (I) and non-infectious disease mortality (N) in total mortality (M=I+N) at t=0. For simplicity, assume that ½ of M is I for group A, while I=0 for group B. Further assume there is no medical progress in N in any period; the trend decline in N mortality is 0 in all periods. Thus for group B, the shock and response are both zero. Accordingly only group A is affected by medical progress.

The progress is entirely confined to I and takes the following form: in period I there is gradual progress that eliminates a fraction $k < 1/2$ of the initial I. The progress accelerates in II when the remaining $(1-k) > 1/2$ of I is eliminated. Thus group A has the following mortality history:

$$\text{Start of period 1: } M_0 = I_0 + N_0$$

$$\text{Start of period 2: } M_1 = I_0 - kI_0 + N_0 \quad (\rightarrow \text{counterfactual trend} = -kI_0)$$

$$\text{End of period 2: } M_2 = N_0 \quad (\rightarrow \text{‘shock’} = -(I_0 - kI_0) - (-kI_0) = (2k - 1)I_0 < 0)$$

$$\text{End of period 3: } M_3 = N_0 \quad (\rightarrow \text{‘response’} = 0 - (-kI_0) = kI_0 > 0)$$

In this example, group A’s (shock, response) pair is a point in the second quadrant; group B’s pair is the origin and the correlation across groups is negative. There is, however, no behavioral significance to this correlation.

⁶ The ex-infectious disease counterfactual implicitly assumes that the discovery of antibiotics would have had no effect on non-infectious disease mortality. However, the model in Dow et al (1999) suggests that there would have been a favorable effect on such mortality. Recall that, in their story, once the infectious diseases have been wiped out, resources would be diverted to reducing other mortality risks. That kind of response implies that the ex-infectious disease counterfactual is too conservative and that the more comprehensive counterfactual used in Table 1 may be the more appropriate.

⁷ Specifically, the comparative static exercises in equations (4) – (6) can be re-worked with the $G(\cdot)$ function assumed, incorrectly, by the chooser to be a constant. The crucial result – a shift to activity Y (the one with non-infectious disease risk) – still goes through.

⁸ Some of the variety can be linked to differences in the importance of infectious diseases. The cross-state, within age-group correlation between the 1938 share of mortality due to infectious disease and the subsequent mortality change is negative for every age group, but mostly smaller than .5 absolutely. There is also some geographic clustering: e.g., Vermont behaves more like Massachusetts than another rural state. So we probably have fewer than 21 true degrees of freedom in this sample.

⁹ A dramatic example is New York. For most age groups it has one of the highest death rates in 1910 and one of the lowest in 1940. In 1910 New York was the focal point of an immigration wave that had dramatically increased population densities (with associated communicable disease effects). By 1940 the population and the doctors had learned how to better cope with urban densities and even to turn them to advantage.

¹⁰ It is roughly consistent with the average of the estimated b 's. However most any alternative that allows substantial convergence gives similar results.

¹¹ The different numbers for the US in Table 6 are due to the use of 5 year age groups here v 10 year groups in Table 2.

¹² This is arguably more appropriate for the rich countries as well as the poor ones. Diffusion of knowledge across countries should cause the rich countries to converge as well as allow the poor countries to catch up. Australia and New Zealand, for example, have relatively modest pre-antibiotic mortality declines. But the plausible reason for this is that they have relatively low mortality at the start of the 20th century. By around 1940 the rest of the rich countries have caught up with Australia and New Zealand. So the sensible counterfactual for those two countries going forward would be similar to any other rich country.

¹³ Closer inspection shows that the continuing progress in poor countries is most evident among the young (under age 35 or so) and the old (over age 60).

¹⁴ Similar but weaker results emerge if the response period is attenuated at 1985. This suggests that any response is also gradual, and it has to overcome the continuing medical progress in the poor countries.

¹⁵ Without this adjustment, the correlations are mainly insignificant.