NBER WORKING PAPER SERIES

ECONOMIC ANALYSIS OF RISK AND UNCERTAINTY INDUCED BY HEALTH SHOCKS: A REVIEW AND EXTENSION

Tomas J. Philipson
George Zanjani

Working Paper 19005
http://www.nber.org/papers/w19005

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2013

This paper is a forthcoming chapter in the NorthHolland Handbook series "Handbook of Economics of Risk and Uncertainty" edited by Mark Machina and Kip Viscusi. We are thankful to the editors for comments that improved the chapter. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Tomas J. Philipson and George Zanjani. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Economic Analysis of Risk and Uncertainty induced by Health Shocks: A Review and Extension
Tomas J. Philipson and George Zanjani
NBER Working Paper No. 19005
April 2013
JEL No. I0,I11

ABSTRACT

We review and extend the economic analysis of risk and uncertainty as it relates to behavior mitigating health shocks. We summarize some central aspects of the vast positive and normative literature on the role of various forms of insurance that attempt to smooth consumption, which can be uneven due to medical spending induced by health shocks. Much of this literature has been concerned with the barriers that prevent full insurance and the role of the government eliminating their adverse consequences. We argue that this large literature is limited in that it is focused largely on consumption smoothing rather than smoothing of health itself. However, a problem with insuring health itself is that human capital cannot be traded; a person diagnosed with an incurable cancer cannot be made whole through reallocation of someone else's health. This lack of tradability in human capital implies that pooling of health risks, through private or public insurance, is infeasible except in rare instances such as transplantations. We argue that medical innovation can be interpreted as an insurance mechanism for a population's health. By enabling treatment of a harmful disease, it completes the previously incomplete market for risk-sharing in health by pooling the health care spending risk. In a sense, medical innovation involves a current certain R&D payment for a reduced future price of health, which is directly comparable to traditional health care insurance where a current premium is paid for a future reduced price of health care. We explore the positive and normative implications of this "health insurance" view of medical R&D and stress the ex ante value of new medical innovations, sometimes for patients that may never even use them. Given the potentially large value of smoothing health itself rather than consumption, we argue that more explicit analysis is needed on the relative value of public programs stimulating medical innovation versus health care reforms largely aimed at enabling consumption-smoothing.

Tomas J. Philipson
Irving B. Harris Graduate School
of Public Policy Studies
University of Chicago
1155 E. 60th Street
Chicago, IL 60637
and NBER
t-philipson@uchicago.edu

George Zanjani
Georgia State University
gzanjani@gsu.edu

Abstract:

We review and extend the economic analysis of risk and uncertainty as it relates to behavior-mitigating health shocks. We summarize some central aspects of the vast positive and normative literature on the role of insurance in smoothing consumption, which can be uneven owing to medical spending as a result of health shocks. Much of this literature has been concerned with the barriers that prevent full insurance and the role of the government in eliminating any adverse consequences due to a lack of insurance. We argue that this large literature is limited in that it is focused largely on consumption smoothing rather than smoothing of health itself. However, a problem with insuring health itself is that human capital cannot be traded; a person diagnosed with an incurable cancer cannot be made whole by reallocating someone else's health. This lack of tradability implies that pooling of health risks through private or public insurance is not feasible except in rare instances such as transplantations. We argue that medical innovation can be interpreted as an insurance mechanism for a population's health. By enabling treatment of a harmful disease, it completes the previously incomplete market for risk-sharing in health by pooling the health care spending risk. In a sense, medical innovation involves a current certain research and development (R&D) payment for a reduced future price of health, which is directly comparable to traditional health care insurance, where a current premium is paid for a future reduced price of health care. We explore the positive and normative implications of this "health insurance" view of medical R&D and stress the ex ante value of new medical innovations, sometimes for patients who may never even use them. Given the potentially large value of smoothing health itself rather than consumption, we argue that more explicit analysis is needed on the relative value of public programs that are designed to stimulate medical innovation versus health care reforms largely aimed at enabling consumption smoothing.

Section 1: Introduction:

Dealing with uncertainties of health shocks has generated much economic and government activity, as well as intense policy debates. Dealing with health shocks has led to a burgeoning insurance market, including life insurance, annuities, workers' compensation, health care insurance, long-term care insurance, and disability insurance, to name a few. A central feature of such insurance is the pooling of financial risks, such that those who are lucky enough to avoid sickness pay for the losses of the unlucky ones who become sick. Accompanying the growth in insurance has been proliferation of economic research on the prevalence, value, and desirability of insurance. In particular, there has been considerable debate on whether public or private financing and production of insurance is more desirable.

In this chapter, we review some of the most commonly discussed results in this vast literature on dealing with health shocks. The goal is not an exhaustive review of various unrelated results, as a complete review of this economic field would indeed be impossible in a single chapter. Rather, the goal is to review the central aspects of this literature, including the often-discussed frictions associated with private health insurance markets, the value of public intervention, and the welfare effects induced by either.

In addition, we call attention to what we regard as a critical gap in the literature on how to efficiently address future health shocks. Specifically, we argue that existing research is focused on insurance against financial shocks and thus does not adequately address a more central and fundamental concern in dealing with health shocks---the restoration of health itself and the smoothing of health across uncertain disease states. Traditional analysis is almost exclusively focused on how to mitigate the impact of health shocks on consumption rather than the impact on health itself. For example, life, annuity, and health care insurance all aim primarily to smooth the consumption uncertainty associated with health shocks. Even when used, however, these arrangements cannot fully

smooth health conditions. To illustrate, when an incurable cancer hits, consumption may be fully insured but what is not covered is the loss in health---and this is the real risk that may impose the largest loss in welfare.

It is important to recognize that because human capital cannot be traded, risk pooling arrangements in health itself, whether through private or public insurance, are often infeasible. For example, if Alzheimer's strikes an individual, he or she cannot be made "whole" or fully healthy by getting health reallocated from someone else.[2] Thus, methods other than risk-pooling must be used to reduce risks to health itself. Given that medical innovation is the primary method by which the real price of health is reduced over time, it can be viewed as serving the role of insuring future health. For example, innovation in treatments for breast cancer and HIV has lowered the price of health, which in turn has smoothed health across such future disease states. Medical innovation is to health what health care insurance is to health care; a certain payment for medical R&D may lower the price of future health while a certain health insurance premium may lower the price of future health care. Thus, medical R&D is "health insurance" in the literal sense of the phrase, as opposed to the colloquial usage where it refers to insurance of health care expenditures. Both are certain investments for uncertain future price reductions and, we argue, may therefore be usefully analyzed and valued using similar methods. In a sense, medical innovation acts like a "financial innovation" that completes a previously incomplete market for health itself by enabling a previously uninsurable shock to be insurable through traditional health care insurance.

While medical innovation is aimed at reducing future risks associated with health shocks, the innovations themselves may involve additional health risks in terms of unsafe side effects. Indeed, governments worldwide extensively regulate the health or safety risks associated with new medical innovations. However, regulations to reduce such health risks affect the type of "health insurance"

[2] A rare exception is when transplantation is feasible, but market mechanisms for such health transfers have been deemed unethical and are outlawed in many countries.

medical innovation involves. They avoid one set of risks, due to side effects of innovations, but impose another, by lowering "health insurance" by negatively affecting innovative returns. We review the literature on the regulation of risks associated with medical innovat*i*ons and suggest new research questions raised by our analysis.

The review unfolds as follows. Section 2 discusses some of the central aspects of the large literature on health care insurance. Section 3 discusses the role of medical innovation as a mechanism to smooth health risk given the infeasibility of pooling such risks. Section 4 reviews the literature on the health risks associated with new innovations, and the public regulation of such risks by the FDA. Section 5 concludes with discussion of useful avenues of future research.

Section 2: Health Care Insurance and Consumption Smoothing

It is well-known theoretical result that a risk-averse consumer prefers full insurance offered on actuarially fair terms under expected utility maximization without state dependence. For example, according to Arrow (1963, p. 961), "… [t]he welfare case for insurance policies of all sorts is overwhelming." Viewed in this light, it is not surprising that state of being uninsured is often taken as *prima facie* evidence of a problem during economic analysis, and economists have devoted much attention to the theoretical and empirical examination of why people are not fully insured. Health insurance is commonly viewed as the canonical example of an inefficient private insurance market, with unusually severe transactional complications that make the presence of the uninsured a pressing policy problem. Indeed, Arrow's endorsement of insurance's value quoted above came in the context of a lament about health insurance.

However, before examining the potential problems in the health insurance market, it is worth noting that, within the world of insurance generally, health insurance take-up is rather high. About 82% of the non-elderly population in the U.S. carried health insurance in 2010.[3] This is comparable to the rate of take-up of automobile liability insurance among motorists (86%) ---which is currently compulsory in most states.[4] Insurance take-up rates from other markets where purchase is voluntary are often substantially lower than that for health insurance among the eligible populations facing the risks involved. For examples, only 11% of California homeowners carry earthquake insurance; 70% of US households carry life insurance; 43% of renters carry renters insurance; and 1% of US properties outside high risk areas carry flood insurance (the take-up rate inside high-risk areas, where purchase is often

[3] "Overview of the Uninsured in the United States: A Summary of the 2011 Current Populaion Survey" *ASPE Issue Brief*, US Department of Health and Human Services.
[4] *Uninsured Motorists* (2011), Insurance Research Council. Cited figure refers to estimates for 2009.

mandated by lenders, is 49%).[5] While each market of course has its own idiosyncratic features, the general point is that health insurance penetration is fairly high when compared with other markets where purchase is not compulsory.

Nevertheless, the canonical model of insurance predicts full insurance under the assumption of actuarially fair pricing, and much economic analysis has been devoted to discussing the causes and consequences of factors disrupting this prediction. Most of this analysis has been under the assumption of the traditional rational choice framework, and our review will largely abide within this framework. It is worth noting, however, that a growing literature exists on the behavioral economics aspects of insurance choice and how relaxing the assumption of consumer rationality affects conclusions about insurance interventions and health policy generally (see, for examples, Abaluck and Gruber (2011); Chernew and Fendrick (2008); Frank (2007); Frank and Lamiraud (2009); Liebman and Zeckhauser (2008); McFadden (2006)).

A natural place to start in tackling the question of why people are not fully insured in a rational choice framework is with what is perhaps the simplest possible explanation--- prices above costs of paying claims, which may be due to production costs in addition to claims (often called "administrative costs" in the context of insurance markets) or markups due to market power. Demand for insurance falls in price as for most other goods, so markups above expected claims costs may reduce or eliminate equilibrium insurance coverage. Arrow's seminal work on insurance contemplated the situation where insurance prices would deviate from actuarially fair values due to production costs of various kinds, such as those relating to agent commissions and to company overhead. The impact on insurance purchase of course

[5] *2011 California Earthquake Premium, Exposure and Policy Count Data Call Summary*, California Department of Insurance; *Trends in Life Insurance Ownership* (2010), LIMRA; 2006 Insurance Research Council survey on renters insurance; Dixon, L., N. Clancy, S. A. Seabury, and A. Overton, 2006, *The National Flood Insurance Program's Market Penetration Rate: Estimates and Policy Implications* (Santa Monica, CA: RAND Corporation).

depends on the nature of the cost structure. Arrow showed that, in circumstances where the premium was determined by the expected loss times a proportional expense loading, the optimal insurance contract would feature full coverage beyond a deductible. Fixed costs could lead to other effects, including non-participation (e.g., Mulligan and Philipson (2004)). As a percentage of premiums, pricing above claims costs are substantial. Such pricing above costs (defined as any cost other than benefit payments to the insured or third party service providers or claimants) can amount to 30% of the premium or more (depending on the line of insurance), illustrating Arrow's (1963) observation that actuarial fairness may not be present in insurance markets generally. Additional markups may be generated by market power. Less work has been done on output restrictions due to market power in the insurance industry, although there is a literature estimating premium elasticities that pertains directly to this issue.[6]

While it is possible that pricing above costs may explain certain patterns of limited coverage and nonparticipation in insurance markets, a lack of universal purchase is hardly unique to insurance products and not by itself necessarily indicative of market failure, even though perceived as such by many economists. Goods and services of course require resources for production. Thus, there is little hand-wringing among economists over production costs unless market imperfections are present. In the case of insurance, contracting costs are emphasized as market imperfections: In particular, moral hazard and adverse selection are often fingered as causes of insurance market failure.

Adverse Selection

[6] For estimates pertaining to individual health insurance, see Krueger and Kuziemko (2013), Strombom, Buchmueller, and Feldstein (2002), and Gruber and Poterba (1994). For a partial survey of estimates in other markets, see Grace, Klein, and Kleindorfer (2004).

Adverse selection commonly refers to a situation of asymmetric information insurers and consumers, where insurers are unable to distinguish the underlying risk characteristics of individual insured. Insured are aware of their own risk characteristics, which of course influence their demand for insurance. The two basic theoretical flavors of adverse selection are the "collapsing market" of Akerlof (1970) and the "separating equilibrium" of Rothschild and Stiglitz (1976).

To illustrate Akerlof's idea with an extreme version (adapted to the context of insurance), the only sustainable market price for contracts is the one corresponding to the expected costs of the worst risks in the consumer population. Any attempt by the insurer to offer a "pooling" contract---in which the price of the contract reflects the average cost in the consumer population---is doomed to failure, as the contract is only attractive to those risks who are "worse than average." As the better risks opt to go without coverage, those remaining in the pool are progressively riskier, so the price must rise---thus continuing the process of driving out the good risks along the margin. The unraveling of the market continues until only the worst risks (the "lemons") are left.

Rothschild and Stiglitz (1976), on the other hand, have a model similar in flavor but less gloomy in terms of predictions. In their model, the insurer still cannot identify individual risk characteristics but can effectively sort consumers by exploiting differences in their desire for coverage. This is accomplished by offering a menu of contracts, with high levels of coverage featuring high per unit prices (as these will be attractive to high risks) and restricted levels of coverage featuring low per unit prices (as these are attractive to low risks but scare off the high risks because of their greater need for coverage).

The economic burden of adverse selection thus falls on the shoulders of good risks, who, in one way or another, end up with less coverage in a second-best world with asymmetric information. And it is this

theoretical observation that has driven much of the subsequent empirical literature: The key empirical test centers on the relationship between coverage and risk: Adverse selection models predict that high risk individuals will purchase more coverage, and it is this connection which lies at the center of most tests.

Viewed across insurance markets---and even within insurance markets---the empirical evidence on adverse selection is mixed (see Cohen and Siegelman (2010) for a cross-market review, some of which is recapitulated here). In life insurance markets, most studies have found no evidence of adverse selection (Cawley and Philipson (1999); Hendel and Lizzeri (2003); McCarthy and Mitchell (2010)).[7] In annuity markets, on the other hand, the weight of the evidence points in the opposite direction, as most studies have found evidence consistent with adverse selection (Finkelstein and Poterba (2002, 2004); McCarthy and Mitchell (2010)). Recent work in auto insurance has tended toward negative findings (Richaudeau (1999); Chiappori and Salanie (2000); Dionne, Gourieroux, and Vanasse (2001); Saito (2006)) although some researchers have found evidence in certain market segments or sublines (Cohen (2005); Muermann and Straka (2012)).

Given the variety of findings in other insurance markets, it is not surprising that the findings in health insurance are similarly mixed. Cutler and Reber (1996) and Altman, Cutler, and Zeckhauser (1998) both find evidence of significant adverse selection in the employer-based health insurance market. Like annuity markets (where adverse selection is also commonly found in empirical studies---see above), employer-based health insurance often features limited or no pricing differentials with respect to the risk characteristics of individual applicants. This suggests that the observed selection effects may owe more to the institutional arrangements in the market rather than the technical ability of the insurer to

[7] He (2009) is an exception, finding a correlation between insurance purchase and subsequent death in HRS data.

gather information. Indeed, Cardon and Hendel (2001) argue that much of the claims variation in health insurance can be explained by observable characteristics, and private information about health status seems to play only a small role.

In other circumstances, asymmetric information between consumers and insurers may be present, but adverse selection---in the sense that the insured population is of higher risk than the uninsured population---is not. This is found by Finkelstein and McGarry (2006) in the market for long-term care insurance and Fang, Keane, and Silverman (2008) in the Medigap insurance market. In these circumstances, it is suggested that preference characteristics other than claims risk are influencing insurance demand. This can actually lead to circumstances of advantageous selection, where lower risk consumers have greater demand for insurance because of other unobservable characteristics. In aggregate, it is possible for the two influences to offset each other---even if there are consumers in the market who are exploiting private information.[8] Thus, asymmetric information is not necessarily *prima facie* evidence of serious inefficiency.

Moral Hazard

Moral hazard was suggested as being significant for the economic analysis of health insurance at least as early as Arrow (1963), who noted that the demand for medical services was influenced by the presence of insurance and that coinsurance provisions in insurance contracts were present to deal with this problem. The idea was extended by Pauly (1968), who clarified how consumer demand for medical services after a health shock varied according to the marginal cost of care---in the extreme case of full insurance; the consumer would face no marginal cost of care and would over-consume care relative to

[8] For a striking example of such exploitation in the long-term care insurance market, see the influence of Huntington's Disease on demand revealed in Oster, Shoulson, Quaid, and Dorsey (2010). Wolfe and Goddeeris (1991) also find adverse selection but conclude that it is unlikely to have serious welfare effects.

the economically optimal level when ill. This behavior was incorporated into the price of the insurance contract, and it was possible to imagine cases where the moral hazard problem would lead even a risk averse consumer to forego insurance because of his or her inability to commit to limiting care utilization after purchase of the contract. Zeckhauser (1970) extended this line of reasoning further by elucidating a main trade-off in insurance contracting connected with cost-sharing provisions. The tradeoff between correct incentives and risk-sharing: cost sharing provisions such as coinsurance could limit moral hazard, but only at the cost of exposing the consumer to greater financial risk.

The theory has been generalized and extended. For example, Goldman and Philipson (2007) analyze moral hazard in the presence of multiple technologies, showing that the main predictions of the single treatment case break down when cross price-elasticities between treatments are nonzero. For example, even though drug demand may be highly elastic, it may be optimal to fully insure it (or even provide subsidies to consume it) to induce less hospital spending in the future. Another direction that seems useful is to extend the theory of insurance under moral hazard to incorporate altruism. This needs to recognize that low co-pays, in addition to raising moral hazard, act as beneficial Pigouvian subsidies from the rich to the poor if the rich care about expanding the care to the poor. The dual role played by copays is exemplified by the low Medicaid copays which according to traditional theory induce excessive moral hazard for the highly elastic care by the poor. This is clearly not inefficient if the rationale for the low co-pays is to stimulate demand of the poor. It seems that the effect of altruistic externalities on optimal insurance design is a useful area of future work.

Moral hazard is difficult to distinguish from adverse selection empirically since both effects work to produce insured populations with higher accident risk (*ex post*) than otherwise similar uninsured (or partially insured) populations. Thus, a positive statistical relationship between coverage and risk is consistent with both moral hazard and adverse selection. Abbring, Chiappori, Heckman, and Pinquet

(2003) show that it is technically possible to distinguish the two in an empirical setting, but panel data is required.

In a health insurance setting, consumer moral hazard could operate either through taking risks relating to health (for examples, lifestyle choices) or through choice regarding the utilization of services once faced with a health shock. The latter effect has been well-documented in the literature, or at least inferred from studies of the price-elasticity of demand for medical services. The RAND Health Insurance Experiment featured the random assignment of families to health plans with different cost sharing provisions. The results clearly indicated that cost sharing provisions reduced overall usage of medical services (Manning, Newhouse, Duan, Keeler, Leibowitz, and Marquis (1987)), and overall price-elasticity of demand in the experiment was estimated to be -0.2---a result echoed in Keeler and Rolph (1988). Though these figures are based on data from the 1970's, more recent work has produced similar findings (Chandra, Gruber, and McKnight (2010)).[9]

Moral hazard of a different ilk is present on the supply side of the market, as care decisions may be heavily influenced by physician recommendations. It follows that medical expenditure, in addition to depending on consumer incentives, will also depend on physician incentives if physician behavior is at all guided by financial incentives. The empirical evidence does indeed suggest that physician behavior is influenced by these incentives (e.g., Gaynor, Rebitzer, and Taylor (2004)).

[9] These estimates are for overall medical expenditure, and much research has focused on components of medical expenditures. To give some sense of the breadth of research available, see Goldman, Joyce, and Zheng (2007) for a survey of price-elasticity of demand for prescription drugs; Kondo, Hoshi, and Okubo (2009) for a study on the elasticity of demand for vaccinations; Connolly, Griesinger, Ledger, and Postma (2009) for research on the elasticity of demand for Assisted Reproductive Technologies (ART).

Given the sensitivity of consumer expenditure to price, it is not surprising that coinsurance, co-payments, and deductibles are common features of health insurance plans. In addition, managed care organizations have introduced numerous cost-saving innovations aimed at the supply side of the market (see McGuire (2012) for a survey). These innovations generally may curb moral hazard and make insurance more attractive to consumers in an *ex ante* sense. Indeed, some might interpret the penetration rate of 82% mentioned above as evidence of the success of the private health insurance market in the US. However, the standards for success in the health insurance market are much higher. Many hold the ideal of universal coverage, and the failure of the private market to adequately address the needs of certain high risk segments of the population---such as the elderly and people with pre-existing conditions---has led to substantial government intervention in the health insurance market in developed countries.

Section 3: Health Stock Insurance and Medical innovation

In this section, we argue that the extensive policy focus on ensuring consumption smoothing in the face of health shocks ignores an important dimension of the welfare effects of such shocks. While it is true that financial shocks relating to health care are significant and deserve attention, much less attention is paid to policies aimed at smoothing health itself in response to shocks. To illustrate, when an individual is diagnosed with Parkinson's Disease, there are two shocks to consider: One is the shock to wealth, since treatments are potentially costly and earnings may suffer; the second is the shock to health itself, since even the most advanced treatments are largely palliative and can only partially restore quality of life. For many people, the second shock may well be the more devastating one from a welfare perspective. Unfortunately, risk-pooling through insurance mechanisms in human capital is infeasible since human capital cannot be traded. This implies that risk reduction from health shocks

must come through other means---the main one being medical innovation. Due to medical innovation, breast cancer and HIV today are shocks to health which are far more "smoothable" than was previously feasible.

More precisely, consider the following simple formalization of the full impact of health shocks and the ex-ante insurance value of medical innovation. Let π_s be the probability of a bad disease state. If healthy, the consumer's health is H_0. If sick, the consumer's health is a function of medical spending, given the technology level x:

$$H_s(M; x) \leq H_0$$

We consider a simple insurance contract with coinsurance $1 - i$ for any medical spending. For example, if the consumer spends M she is reimbursed $i * M$ by insurance, with $0 \leq i \leq 1$. The premium is assumed to reflect the actuarial value times a loading factor $L(\tau) \geq 1$:

$$P(i, M) = (\pi_s * i * M)L(\tau)$$

where we assume that τ represents the technological efficiency of the insurance market so that $L' < 0$.

To focus on the central issues here, we assume away moral hazard. The consumer thus chooses both the coinsurance rate and medical care ex ante to maximize

$$\mathrm{V}(\mathrm{Y}, \mathrm{x}, \tau) = max_{M,i}\ \{\pi_s U\big(Y - (1-i)M - P(i,M), H_s(M;x)\big) + (1-\pi_s)\, U(Y - P(i,M), H_0)\}$$

The FOC for optimal ex-ante care leads to:

$$\left(\pi_s \frac{\partial U_s}{\partial H_s}\right)\frac{\partial H_s}{\partial M} = \pi_s \frac{\partial U_s}{\partial C}(1-i) + \left(\pi_s \frac{\partial U_s}{\partial C} + (1-\pi_s)\frac{\partial U_0}{\partial C}\right)P'$$

with the first term on the right hand side representing the marginal cost associated with medical care copayment and the second term representing the marginal cost associated with premiums.

There are two key interrelated risks faced by the individual. The first is consumption risk, a risk which is created by uncertainty in the level of medical care spending. The second is health risk, which is

mitigated by the utilization of medical care. The vast majority of the literature dealing with impacts of "health care insurance" is concerned with the former risk.

Improvements in the technological efficiency of the medical insurance market (represented in our simple model as increases in τ) enable better smoothing of consumption in the presence of health shocks and may improve health as well if optimal medical care spending increases in response to decreases in the effective price of care:

$$\frac{dH_s^*}{d\tau} = \frac{\partial H_s^*}{\partial M}\frac{\partial M^*}{\partial \tau} > 0$$

with an ultimate impact on welfare composed of an impact on consumption and an impact on health as in:

$$\frac{\partial \mathrm{V}}{\partial \tau} = \left[\pi_s \frac{\partial U_s}{\partial C}\frac{dC_s^*}{\partial \tau} + (1-\pi_s)\frac{\partial U_0}{\partial C}\frac{dC_0^*}{\partial \tau}\right] + \pi_s \frac{\partial U_s}{\partial H}\frac{dH_s^*}{\partial \tau}$$

Through its influence on both consumption and health, the change in technology affects both consumption risk and health risk.

The second risk--- risk to health itself---is less often addressed. As noted above, it is possible that improvements in the efficiency of insurance will improve health, and this has been analyzed extensively (for a review see Levy and Meltzer (2008)). However, the most important limitations on the health risk mitigation will often concern medical technology. As it is impossible to have risk-pooling for human capital that is not tradable, limiting health risk will have to come from other means. The main one is medical innovation as represented by an increase in the productivity of producing health from medical care, the parameter x above. If we interpret increases in x as representing improvements in the productivity of medical care spending (i.e., $\frac{\partial^2 H_s}{\partial M \partial x} > 0$), then the first order effect of increases in medical technology will be improvements in health. More productive health care raises the marginal product of care, and thereby raises the optimal health stock upon a health shock:

$$\frac{dH_s^*}{dx} = \frac{\partial H_s^*}{\partial M}\frac{\partial M^*}{\partial x} + \frac{\partial H_s^*}{\partial x}$$

and, more to the point, increases consumer welfare:

$$\frac{\partial \mathrm{V}}{\partial \mathrm{x}} = \left[\pi_s \frac{\partial U_s}{\partial C}\frac{dC_s^*}{dx} + (1-\pi_s)\frac{\partial U_0}{\partial C}\frac{dC_0^*}{dx}\right] + \pi_s \frac{\partial U_s}{\partial H}\frac{dH_s^*}{dx}$$

This simple analysis raises the key question of the relative value of improvements in medical technology (x) versus improvements in insurance (τ). At the margin, the consumer will value improvements in medical technology more than improvements in health insurance efficiency if:

$$\frac{\partial \mathrm{V}}{\partial \mathrm{x}} > \frac{\partial \mathrm{V}}{\partial \tau}$$

The value of improvements in medical technology to the consumer may often be much larger than the value of consumption smoothing. To illustrate, consider an incurable disease for which the current medical technology x is ineffective. In this extreme case:

$$\frac{\partial H_s}{\partial M} = 0,$$

Nothing is spent on care as it is unproductive, so expected utility is

$$\mathrm{V}(\mathrm{Y}, \mathrm{x}, \tau) = sU\big(Y, H_s(0;x)\big) + (1-s)\,U(Y, H_0)$$

Even though there is perfect consumption smoothing there is of course a loss in health induced by the health shock, and the value of traditional health care insurance is obviously zero. Moreover, gains in insurance market efficiency obviously yield no gains in welfare:

$$\frac{\partial \mathrm{V}}{\partial \tau} = 0$$

since medical spending is unproductive under the current state of technology. Even if insurance were free, it would have no value: Gains in welfare in this scenario can come only from reducing the price of health through medical innovation.

Now consider when medical innovation progresses in a way that more health can be obtained for successively lower amounts of medical care spending. This implies in the extreme there is no loss in

health and minimal resources used to restore health upon a health shock. In this case, expected utility converges to $U(Y, H_0)$. In this case, with the most extreme form of perfect medical productivity, it is as if no health shock occurred in the first place. In this example, gains in medical innovation will eventually not only insure health smoothing but eliminate the need for consumption smoothing related to medical expenditure as medical technology reduces the cost of care. Put differently, there is no need for insuring consumption when faced with cheap care.

The value of being relieved of health risk can be defined by the willingness to pay W to transition between the two extreme forms of medical care productivity (completely unproductive to perfectly productive). In other words, how much income would one be willing to sacrifice in the high productivity world (where health risk is absent) to make one indifferent to the low productivity world, where the individual is still subject to health risk:

$$U(Y - W, H_0) = sU\big(Y, H_s(0; x)\big) + (1 - s)\, U(Y, H_0)$$

It measures how much money one would be willing to pay when healthy (high productivity case) to avoid facing the risk of disease (low productivity case), and is clearly greater than zero in this case---which is the value of smoothing consumption in the low productivity world.

More generally, it is important to note that the value of consumption smoothing evidently depends on the state of medical technology. In the extreme illustration above, there is no value to consumption smoothing when medical technology is completely ineffective. In addition, there also is no value when the medical technology is so effective that only minimal medical expenditure is required to restore health. However, intermediate states of medical technology, where improvements in health can be purchased at non-negligible cost, will be associated with gains to insuring medical expenditures.

The reverse is also true: The value of medical technology improvements also will, in general, depend on the efficiency of the insurance market as represented by τ. As will be illustrated more

formally in Section 3.2, access to an expensive medical technology may be possible only with insurance, meaning that the value of the technology depends on the existence of the insurance market.

3.1 Medical R&D as Insurance

Given the dual impact of health shocks on health and consumption, medical innovation serves the role of reducing adverse health events in future disease states and thus may be valued in a similar manner to other forms of risk reduction. Consider when there are two technologies that may be used to treat a disease, denoted by x_1 and x_2. The health outputs associated with these two technologies follow a Leontief production function requiring medical care inputs of M_1 and M_2 respectively. That is, for $k = 1,2$:

$$H_s(M; x_k) = H_s(0; x_k) < H_s(M_k; x_k)$$

The second technology is more effective in restoring health, so that:

$$H_s(M_1; x_1) < H_s(M_2; x_2) \leq H_0$$

but has not been developed. Suppose initially that R&D investment of size R per person is required to develop the technology and is certain to succeed. In this case, the utility associated with undertaking the research and development effort (assuming the new technology is preferred to the previous one when developed) is:

$$\mathrm{V}(\mathrm{Y} - \mathrm{R}, x_2, \tau) = max_i \{\pi_s U\big(Y - (1-i)M_2 - P(i, M_2) - R, H_s(M_2; x_2)\big) + (1-\pi_s)\, U(Y - P(i, M_2) - R, H_0)\}$$

while utility under the previous technology was

$$\mathrm{V}(\mathrm{Y}, x_1, \tau) = max_i \{\pi_s U\big(Y - (1-i)M_1 - P(i, M_1), H_s(M_1; x_1)\big) + (1-\pi_s)\, U(Y - P(i, M_1), H_0)\}$$

Notice the parallel between the R&D effort and insurance. The R&D "premium" is simply the per person development cost R which is paid up front. The benefit is an in-kind "claim payment" in the form of improved health in the event that the consumer becomes sick. Although we have not modeled

it as such, the claim payment could be uncertain in that the R&D effort might have uncertain prospects for success, but the overall impact is an expected improvement in health status in the sick state.

One can thus interpret R&D as a form of insurance of health itself---in which a payment of a development cost premium is made in exchange for a lower price of improved health in case of illness. Thus, medical R&D is "health insurance" in the literal sense of the phrase---as opposed to the colloquial usage where "health insurance" refers to insurance of medical expenditures.

This distinction between insuring health and insuring medical expenditures brings up an important tradeoff between policies aimed at enabling better health and those aimed at consumption smoothing. Public resources can be spent on health insurance subsidies for *existing* technologies (for example, by granting tax breaks for health insurance purchase, or through public provision of existing medical technologies) or on R&D investment to generate *new* technologies. Medical expenditure insurance subsidies under the existing technology can be represented here as an increase in insurance market efficiency from τ to τ', with the total spent on the subsidy being:

$$[L(\tau) - L(\tau')] * \pi_s * M_1 * i^*(Y, \tau', x_1)$$

where $i^*(Y, \tau', x_1)$ is the co-insurance chosen by the consumer in the presence of the subsidy. To compare the relative efficiency of public subsidization of consumption smoothing versus health smoothing, we may consider τ' chosen so that amount spent on the expenditure subsidy is the same as that spent on R&D investment:

$$[L(\tau) - L(\tau')] * \pi_s * M_1 * i^*(Y, \tau', x_1) = R$$

The value of the expenditure subsidy in terms of willingness to pay can be defined as:

$$V(Y - W^E, \tau', x_1) = V(Y, \tau, x_1)$$

while the value of the R&D investment can similarly be defined as:

$$V(Y - W^R, \tau, x_2) = V(Y, \tau, x_1)$$

Thus, the value of buying the (perhaps uncertain) health restorations through medical R&D is larger than the value of buying consumption smoothing with health insurance subsidies whenever:

$$W^R > W^E$$

An important component of the value of buying future reductions through medical R&D or health insurance is the value of health improvements from existing and future technologies. There is a large literature on the "value of life" or health more generally that bears directly on this issue. Viscusi (2003) reviews the large literature on this issue. Murphy and Topel (2006), using similar methods, estimate that improvements from a 1% reduction in mortality from cancer would be worth about $500 billion a year. It seems as an open question for future research comparing empirically such estimates of the value of medical R&D to the relative value of consumption smoothing from health insurance.

3.2 Interactions between health care insurance and health insurance through medical R&D.

The foregoing analysis illustrates the potential value of "health insurance" provided by medical innovation as opposed to the consumption smoothing via traditional health care insurance. In this section we discuss we discuss existing papers and suggest new work on how the two types insurance, medical innovation and health care insurance, interact in terms of how one affects the incentives for and value of the other.

3.2.1 Health care insurance and medical R&D incentives

As recognized as early as the patent clauses of the US Constitution, R&D in general and medical innovation in particular needs to be supported by profits and adequate pricing. Weisbrod (1991) stressed the relationship between health care insurance or third-party pricing for the type of medical

R&D undertaken and stressed the role of quality enhancing R&D in lack of reimbursements that incentivize cost reductions.

Subsequent empirical work has found support for this idea. Pauly and Danzon (2002), argue that the rise of prescription drug coverage is likely to have spurred investment in R&D. Finkelstein (2004), Finkelstein (2007)), and Clemens (2012) have documented a positive impact of insurance coverage on medical innovation. These papers stress the dynamic impact of coverage beyond the positive static incentive effects on utilization from lower demand prices or co-pays.

Hult and Philipson (2012)," analyze explicitly how public insurance reforms affect the returns from medical innovation. Pioneered by the work of Newhouse (1992), research suggests that medical innovation is central to the growth in health care spending (see also Chernew and Newhouse (Handbook of Health Econ 2011, ch 1)). Moreover, public reforms are central to driving global innovative returns, as a large share of the world's care is publicly financed in rich countries. Therefore, public reforms have large effects on the uncertain future profits associated with medical innovation, which in turn drive spending growth in both the public and the private sector. The analysis considers cases in which the impact of government reforms on medical research and development (R&D) returns comes from three different sources: expected cash flows, the timing of the flows, and the risk adjustment of those cash flows. For the impact on expected cash flows, the analysis stresses the non-monotonic effects of government expansions on innovative returns. In particular, government expansions often lower both demand prices (copays) and supply prices (reimbursements) through government monopsony power. This may imply that R&D returns rise when government expansions include poorer parts of the population by raising quantity more than lowering markups. For example, the recent Medicaid expansions of Affordable Care Act in 2010 raise innovative returns in this manner. However, innovative returns fall when public insurance expansions include richer parts of the population if markups go down more than utilization goes up. For example, the single-payer European payment systems lower

innovative returns in this manner. The non-monotonic impact of government expansions across the income distribution implies that government cutbacks may raise R&D returns, and pose upward pressure on future public liabilities. Likewise, government expansions may lower public liabilities.

Related to how reform affects innovative returns, Koijen, Philipson, and Uhlig (2012) documented a large "medical innovation premium" that historically is paid to investors and the growth of the health care sector this premium implied. The paper provides an explicit analysis of the link between financial- and real markets for health care by considering how the returns to medical R&D interact with the growth of the sector. The paper documents evidence of a “medical innovation premium,” a large risk premium of about 4-6% annually higher than is predicted by benchmark asset pricing models for firms engaged in medical R&D. They interpret this premium as compensating investors for bearing risk with respect to public health insurance reforms, and the paper analyzes its quantitative implications for the growth of future health care spending. The calibration implies substantial effects of the premium on innovation and health care spending, on the order of magnitude of 4% of GDP, and therefore is argued to be important for future projections of the size of this sector.

3.2.2 Medical R&D for Rare Diseases

Many countries disproportionally subsidize medical innovation for rare or orphan diseases. For example, the Orphan Drug Act in the US provides R&D subsidies for diseases affecting less than 200,000 patients. Even private payers, moreover, often pay very high per-capita reimbursement rates under the rationale that the smaller the disease the larger does per-capita revenues have to be in order to support a given R&D investment. This behavior seems puzzling since the same R&D spending could be used to help a larger set of patients; a small market size should lower efficient R&D according to traditional efficiency arguments. Indeed, a number of orphan drugs feature annual price tags well over $100,000---

and, thus, well beyond the means of the typical patient.[10] In the context of the usual model of financial insurance, it is hard to understand how this---the insurance of losses in excess of wealth, as well as investment in treatments that the affected consumers cannot afford---could possibly be efficient.

However, with the ex-ante insurance role of medical innovation, rare disease R&D may well be efficient. This is because the R&D is essentially acting as an insurance against a low probability event that may often involve severe reductions to the health stock. Such insurance is similar to life insurance used to smooth consumption of beneficiaries: In both cases, the smoothing target is not simply the wealth or consumption of the purchaser. In the case of rare diseases, ex-post per-capita pricing of treatments that are inversely related to prevalence to support R&D may be efficient when considering their value in terms of smoothing health. .

The presence of risks to one's health stock, as distinguished from financial risk relating to medical care expenditures, complicates the analysis of optimal insurance of medical care expenditures by introducing a form of state dependence. Specifically, the marginal value of financial wealth can conceivably change dramatically according to whether one is in the "sick state" or the "healthy state", which can lead the optimal insurance contract to feature what appears to be "over insurance" or "under insurance" of financial shocks associated with the sick state as a consequence of state-dependent utility (e.g., Cook and Graham (1976); Dionne (1982); Nyman (1999)). This could in principle lead consumers either to transfer wealth into the sick state (Zeckhauser (1973)) or to underinsure the sick state (see Pauly (1990)). Viewed in this light, the optimal transfer of wealth into a low-probability sick state could in principle be extreme---resulting in the expenditure of resources far beyond an individual's wealth in the sick state. Importantly, these effects may justify investment in very expensive treatments.

[10] "The World's Most Expensive Drugs" by Matthew Herper, Forbes.com 2/22/2010, accessed on 12/2/2012.

To illustrate, consider a case of a cure for an otherwise untreatable rare disease. Suppose the cure is costless to implement once developed but has a development cost of R (per diagnosis) that greatly exceeds typical wealth levels

$$R \gg Y$$

This development cost must be recovered from each new diagnosis. This would seem at first glance to be an unviable treatment since the consumer cannot afford to pay for the cure. However, with a medical expenditure insurance market, the consumer would be able to buy insurance priced at

$$P = \pi_s * R * i * L(\tau)$$

And will do so if

$$\pi_s U(Y - (1-i)R - P, H_0) + (1-\pi_s)U(Y - P, H_0) > \pi_s U(Y, H_s) + (1-\pi_s)U(Y, H_0)$$

Insurance in this case ensures "access" to an expensive treatment, an interpretation of the over-insurance of the sick state stressed by Nyman (1999). Insurance also enables *development* of the treatment, which would not otherwise be possible. Thus, medical expenditure insurance and R&D "insurance" can be complementary in the sense that the presence of one increases the demand for the other and vice versa.

Importantly, in this case the complementarity is efficiency-enhancing. From a health shock perspective, rare disease R&D may be efficient because it provides insurance against a small probability, but severe, health shock. In other words, rare disease R&D is a fixed payment today to potentially have a restoration of health in the case the small probability shock occurs. Even though there is a small market for the rare disease product once marketed, the *difference* in health across the two states is reduced ex-ante by the medical R&D, and the ex ante effect on welfare may more than justify the expenditure of the "premium."

The foregoing argument is predicated on the notion that individual valuation of life may exceed financial resources. Yet the literature on the issue seems to support this. Existing estimates of the value of a statistical life (VSL) produce a wide range of answers, but a typical answer is well into the millions of dollars (see Viscusi, Woock, and Ziliak (2012) for a recent survey of the challenges in this literature and new estimates in the 4-10 million dollar range). Estimates are typically based on observed willingness-to-pay for mortality risk reductions (or willingness-to-accept risk increases). Estimates of the VSL well into the millions, however, are paradoxical in the sense that the value of life is put far beyond the resources (e.g., the discounted present value of labor income) of the median individual. Such a finding, however, is consistent with the notion of health stock smoothing. As argued above, the willingness to pay for a mortality risk reduction in a fatal disease from a small probability π_s to zero (i.e., through some medical innovation) could well exceed $\pi_s * Y$. Further suggestive evidence of the willingness of individuals to commit extreme resources to the preservation of health can be found in bankruptcy statistics. Recent evidence (Himmelstein, Thorne, Warren and Woolhandler (2009)) suggests that medical expenses are the major cause of more than half of all personal bankruptcies.

Section 4: Health Risks of Treatments, Their Regulation, and the Impact on Health Smoothing through Medical innovation

The discussion so far stressed the ex-ante health insurance role of medical innovation. However, new innovations and treatments may introduce new risks to health in themselves, through unsafe products with side effects. In virtually all developed countries and many developing countries, governments provide regulatory oversight over the health risks of products generated by medical innovation. In the United States, this oversight is conducted by the Food and Drug Administration (FDA), which regulates drugs, medical devices, biologics (products made from living organisms, like vaccines

and blood products), cosmetics, radiation-emitting electronic products, veterinary products, and food. According to the FDA, the products it regulates account for more than one-fifth of U.S. consumer spending in 2010.

The manner in which the FDA regulates the quality or health risks of medical products has a substantial impact on the cost of their development and thus on the speed at which medical innovation can ensure health smoothing when health shocks occur. The FDA requires that companies conduct clinical trials to demonstrate that their medical products are safe and effective. These trials account for a large portion of the total development costs of these products (DiMasi, Hansen, and Grabowski, 2003; Adams and Brantner, 2006). In addition, completion of trials does not guarantee that a product will be approved. This risk of non-approval compounds the cost of product development (DiMasi, Hansen, and Grabowski, 2003).

Despite the central role of the FDA in regulating the quality and R&D costs of medical products, there has been relatively little theoretical or empirical research conducted by economists on the efficiency of FDA policies, particularly as they relate to the ex ante insurance role of medical innovation. Ironically, if a product application was presented to the FDA with the same scant amount of evidence that currently exists on the efficiency of the policies of the agency itself, such an application would likely be rejected on the basis of insufficient evidence.

Despite the lack of work on the ex-ante or ex-post efficiency of FDA policies, a substantial literature has emerged on descriptive aspects as well as the effects of various policies (See Malani and Philipson (2012)). The FDA aims to economize on transaction costs in verifying product quality, a verification that would be very difficult for an average citizen to do. Early static analysis of FDA policies, starting with the papers by Wardell (1973) and Peltzman (1973a), however, have raised concerns about the impact of FDA regulation on the expected profits of medical product companies and thus their incentive to innovate. Moreover, FDA regulation surely increases the cost of R&D by requiring the

generation of extra information through costly clinical trials. Together these factors would reduce the return to, and thus the amount of, R&D investments intended to generate new medical products.

An important aspect of regulating the health risks of medical products is that in many countries including the United States, medical products are jointly regulated by agencies such as the FDA, which screens products to ensure they are safe and effective before they are sold, and the tort liability system, which allows patients to sue manufacturers after they have consumed these products. This "dual" aspect of product safety regulations for health risks has been analyzed by Philipson, Sun, and Goldman (2011) who argues that one form of regulation may increase costs in the presence of the other. Work by Philipson et al. (2008) considers the dynamic welfare effects of FDA regulation by considering the present value of all future costs and benefits of the products being regulated.

Most of the literature on the FDA has to date been descriptive empirical analysis or analysis estimating the effects of various interventions. Empirical analysis of FDA regulation can be grouped into at least five categories

- The number of chemical entities introduced,
- Development costs,
- Development and review times,
- Withdrawal rates, and
- Demand and supply curves for drugs in order to measure changes in consumer and producer surplus.[11]

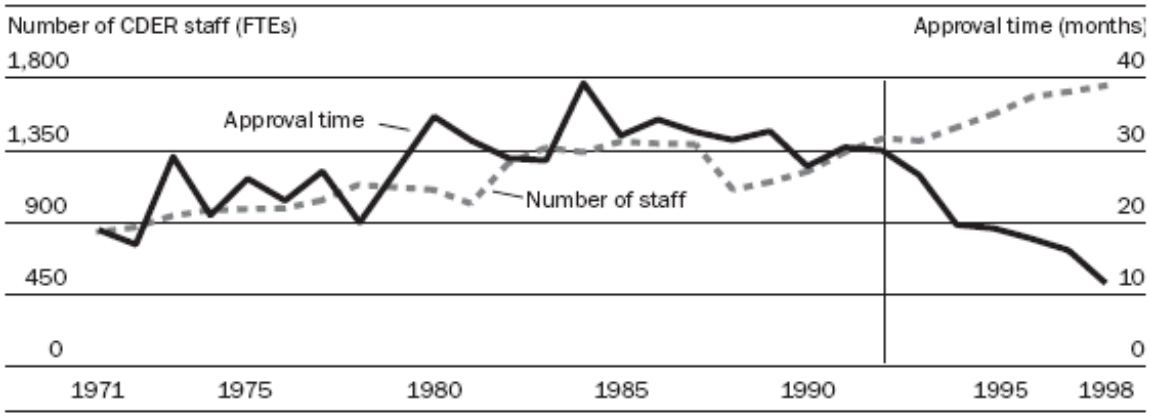
There are two basic challenges to identifying how FDA regulation affects medical innovation and development. One is how to "quantify" FDA regulation. Researchers have taken two basic approaches.

[11] Specifically, a proper welfare calculation requires separately estimating the lost surplus from products that are not approved by the FDA due to minimum quality regulations and the demand curve for products had the FDA not provided more accurate information on quality.

One is to look at adoption of any pre-market clearance regulation, such as the 1962 amendments in the US (e.g., Peltzman, 1973a). This treatment is coded as a dummy variable, set to 0 before 1962, and 1 after.[12] The other is to proxy for regulation by the time it takes for the FDA to review a new drug application (NDA) (e.g., Wiggins, 1981; Jensen, 1987; Berndt, et al., 2005a, 2005b; Carpenter, Zucker, and Avorn, 2008a; Philipson, et al., 2008). This has varied substantially over time. In 1960, approval times were roughly 5 months. After the 1962 amendments, approval times rose dramatically, reaching 20 months in 1970. For most of the 1980s approval times hovered between 30 to 35 months. Approval times declined substantially after the passage of PDUFA in 1992. By 1998, approval times were approximately 12 months, which is roughly where they stand today. This rise and fall in approval times is illustrated in the Figure 1 below, which is reproduced from Olson (2004).

Figure 1

Food And Drug Administration (FDA) Center For Drug Evaluation And Research (CDER) Staffing Levels And New Molecular Entity (NME) Approval Times, Averaged By Year Of Approval, 1971–1998



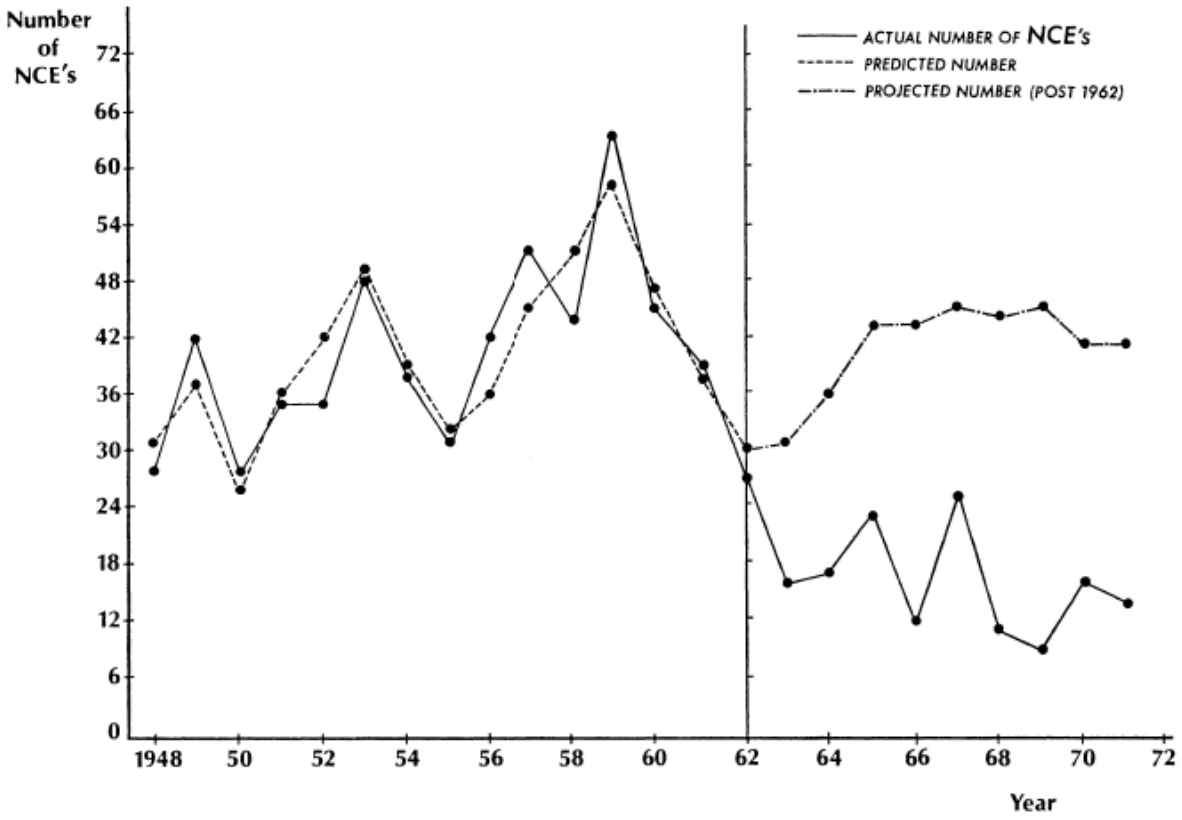
SOURCES: Number of full-time-equivalent (FTE) CDER staff (dashed line) is from the Food and Drug Administration. Approval times (solid line) are from Pharmaceutical Research and Manufacturers of America (PhRMA), based on FDA reports.
NOTES: Data on approval times for 1979 were not available; the trend between 1978 and 1980 was provided by linear interpolation. The vertical line denotes the year in which the Prescription Drug User Fee Act (PDUFA) was passed, 1992.

The second challenge is constructing a baseline against which to judge the effect of the FDA regulation. Because the FD&C Act is a national statute, researchers cannot use, for example, differences in

[12] Similarly, studies that examine the UK code the treatment dummy as 0 before 1973, when the UK adopted pre-market screening for efficacy, and 1 after that (Grabowski, Vernon, and Thomas, 1978).

outcomes across US states that regulate drugs and states that do not regulate drugs. This makes it difficult to separate effects of the statute from underlying time trends. Researchers have used two basic methods to overcome this problem. One is to assume a parametric structure for outcomes in the absence of the 1962 amendments. This could be as simple as including a time trend in the regression. Or it could involve something more elaborate. For example, Peltzman (1973a) used pre-1962 data to estimate a model of new drug introductions and then predicted baseline new drug introductions after the amendments by inserting post-1962 data into his estimated model. When he plotted actual introductions of new chemical entities against his predicted introductions, the result was a striking plot that became popular among critics of the FDA. His figure is reproduced below as Figure 2:

Figure 2 – Actual versus Predicted New Chemical Entities, 1948-1972



The other approach researchers have used to construct a baseline is to examine the development of drug markets in countries that are similar to the US but either did not pass strict drug

regulation in 1962 or took less time to review new drug applications. The primary candidate is the United Kingdom, which passed pre-market clearance for safety in 1963 but did not require proof of efficacy before sale until 1971 (Grabowski, Vernon, and Thomas, 1978). The UK also had shorter review times than the US, at least until the passage of PDUFA in 2002. For example, in 1980, total development times (including preclinical testing, clinical testing and regulatory review times) were 145 months in the US versus just 70 months in the UK (Thomas, 1990).

Tables 1 – 3 summarize research on the effects of FDA regulation on three important sets of outcomes. Table 1 focuses on innovation and includes outcomes such as new drug introductions and the productivity of R&D expenditure (i.e., new drug introductions/R&D expenditures). Table 2 examines drug development and FDA approval times. Table 3 considers the effect of FDA regulations on safety. The main outcomes are involuntary drug withdrawals. The tables not only report findings, but also the data employed, how FDA regulation is measured (e.g., 1962 dummy or review time), and how the counterfactual or baseline is constructed (e.g., parametric time trend or international comparison).

Table 1. Review of literature concerning the effect of FDA regulation on innovation.

Source	Data (usually the dependent variable)	Measure of FDA regulation	Baseline/control	Finding
Wardell (1973)	NCE introductions, 1962-1971	1962 Amendments	UK	Annual NCE flow falls 54% due to 1962 amendments.
Peltzman (1973)	NCE introductions, 1948-1971	1962 Amendments	Model of NCEs using pre-1962 data	Annual NCE flow falls 66% due to 1962 amendments.
Grabowski, Vernon & Thomas (1978)	NCE flow/R&D expenditures, 1960-1974	1962 Amendments and NDA approval times	UK	1962 amendments increased avg. cost of NCE by factor of 2.3 (using 1962 dummy) or 1.9 (using approval times).
Wiggins (1981)	NCE introductions, 1970-1976	NDA approval times	Therapeutic classes with shorter approval times	Increase in approval times due to 1962 amendments decrease NCE introductions 52%, holding R&D expenditures constant; accounting for effects of longer approval times on expenditures, reducing delay to pre-1962 levels would increase NCE introductions by 135%.
Wiggins (1983)	R&D expenses	NDA approval	Therapeutic	Approval times reduced R&D

	(from PHRMA) by therapeutic class, 1965-1968, 1971-1976	times	classes with shorter approval times	expenditures during 1971-1976, but not 1965-1968, possibly because it took time for drug companies to determine how stringent FDA regulation would be after 1962.
May, Wardell & Lasagna (1983)	Number of NCEs tested on humans and NDA approvals, 1958-1979	1962 Amendments	Pre-1962 period	NCEs tested on humans fell from 89/year to 17/year in 1979; NDA approvals fell by 49%.
Cullen (1983)	190 drug product launches across 18 countries during 1961-1976	Surveyed 6 companies for their views of "regulatory tightness" in different counties in 1982. Ratings from 1 (most stringent) to 5 (least stringent).	17 countries other than US	Countries rated as having tighter regulations had (1) a larger increase in lag between first introduction in any country and introduction in that country from the 1960s to the 1970s and (2) a smaller increase in the number of products introduced in that country from 1960s to 1970s.
Jensen (1987)	NCE introductions by 26 firms 1969-1979	NDA approval times	Classes with shorter approval times, time trend	One month decrease in approval times increase annual NCE introductions by 15%.
Thomas (1990)	NCE, sales and market cap of drug companies, 1960-1980	1962 Amendments and approval times	UK	FDA regulation did not affect NCEs at large firms, but did substantially reduce NCE introduced by small firms. Due to reduced competition from small firms, sales rose at large firms in the US.

Table 2. Review of literature concerning the effect of FDA regulation on approval times.

Source	Data (usually the dependent variable)	Measure of FDA regulation (or other treatment variable)	Baseline/ control	Finding
OTA (1989)	Effective patent length of drugs	Ex post commercial importance of drug		Drugs with greater ex post commercial importance have longer effective patent length.
Thomas (1990)	Preclinical testing, clinical testing, and NDA review times, 1960-1980	1962 Amendments	UK	US total development times grew from 35 months in 1960 to 120 months in 1970 to 145 months in 1980. The increases in preclinical testing, clinical, and NDA review times were 30, 60 and 20 months, respectively. In the UK, total development times increased from 30 to 70. Preclinical testing times were constant while the sum of clinical testing and review times increased by 40 months.
Kaitin et al. (1991)	Approval times	FDA ratings novelty of drugs		FDA accelerated approval of more novel chemical entities.
Dranove & Meltzer (1994)	Time from drug patent application to NDA approval for 564NMEs between 1950-1986	Various measures of importance of drug (e.g, FDA rating, commercial value, citations, worldwide introductions)		Development and approval times are lower for more important drugs.
Carpenter et al. (2003)	Approval times and FDA (CDER) staff, 1971-1998	PDUFA	Time trend	Funding for FDA staff has bigger influence on NDA review time than source of funding (user fees under PDFUA).
Olson (2004)	Approval times and FDA (CDER) staff, 1971-1998	PDUFA	Time trend	PDUFA reduced approval times by 34% by 1998. Different result than Carpenter et al. (2003) because Olson groups approvals by approval year rather than NDA submission year as Carpenter et al do.
Berndt et al. (2005a)	Clinical development and NDA review times, 1965-2003	PDUFA	Time trend	PDUFA reduced approval times by 7.6% per year during PDUFA I (1992-1996), and 3.6% per year during PDUFA II (1997-2001). PDUFA II may also have reduced clinical development times by 4.5%.

Table 3. Review of literature concerning the effect of FDA regulation on safety.

Source	Data (usually the dependent variable)	Measure of FDA regulation	Baseline/control	Finding
Bakke, Wardell & Lasagna (1984)	Drug discontinuations, 1964-1983	1962 Amendments	UK	Few discontinuation in either country so no significant differences in discontinuations in US vs. UK
Bakke et al. (1995)	Drug discontinuations, 1974-1993	1962 Amendments	UK, Spain	More drugs discontinued in UK (20) and Spain (16) than US (10). Normalizing by number of drugs approved shrinks the difference: 4% in UK vs. 3% in US.
GAO (2002)	Drug withdrawals, 1986-2000	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals were 3.1% in 1986-1992 and 3.5% in 1993-2000.
CDER (2004)	Drug withdrawals, 1971-2004	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals were 2.7% in 1971-1993, 2.3% in 1994-Apr. 2004.
Berndt et al. (Nature, 2005b)	Drug or biologic withdrawals, 1980-2000	PDUFA	None	No significant effects of PDUFA on withdrawals. Withdrawals (including biologics) were 2.8% in 1980-1992, and 2.2% in 1993-2000.
Carpenter, Zucker & Avorn (2008)	FDA withdrawals, black-box warnings and voluntary withdrawals by drug companies, 1993-2004	PDUFA	Drugs approved well before or after PDUFA deadlines	PDUFA caused bunching of FDA approval during 2 months before deadlines. Drugs approved in 2 months before deadlines had higher odds of being withdrawn by the FDA (OR = 5.5), getting blackbox warnings (4.4) and of being voluntarily withdrawn (3.3) than drugs approved well before or after deadlines

4.1 Innovation

The initial papers studying the effect of FDA regulation on innovation used the 1962 amendments as a treatment and the number of new chemical entities (NCE) introduced each year as the outcome. Whether they used the UK (Wardell, 1973) or a model of introductions fitted to pre-1962 data (Peltzman, 1973a) as the controls, they found large reductions in NCE introductions associated with the legislation. The chart from Peltzman (1973a), reproduced above in Figure 2, is illustrative.

The Peltzman paper was criticized, however, for overestimating the reduction in NCEs.[13] First, it examined only the quantity of drugs approved, not their quality. Perhaps only relatively unimportant drugs were held back in the 1960s. Second, drug companies may have voluntarily reduced NCE introductions even without the 1962 amendments. They may have interpreted the Thalidomide controversy as evidence of increased consumer demand for safety and stopped developing drugs that had substantial side effects. Coupled with the great advance in the ability of the pharmacological sciences to detect side effects from drugs, companies may have held back drugs for fear of losing good will or facing legal liability. Third, given the high value of drugs developed in the 1950s and 1960s, it is possible the returns to drug development had simply diminished by the 1960s (Grabowski, Vernon and Thomas 1978).

A second round (Grabowski, Vernon, and Thomas, 1978; Cullen, 1983; Thomas, 1990) of papers therefore focused on the UK as a control for the US. The UK experienced the same increase in demand for safety after the Thalidomide controversy and potentially diminishing returns in drug development. Yet the UK only introduced pre-market testing for safety in 1963, and did not introduce testing for efficacy until 1973. Therefore, comparing the US and UK in the 1960s would highlight the effect of pre-

[13] Wardell's papers, e.g., Wardell (1973), were widely cited but did not receive serious attention in the economics literature. This may be because the papers did not employ any serious statistical analysis to probe the findings.

market screening for efficacy. These UK comparisons also showed significant reductions in research output associated with the increased US regulation.

One problem with studies that focused on the 1960s, according to Wiggins (1983), was that it took some time for the FDA to decide how to implement the 1962 amendments. Moreover, it also took drug companies some time to learn how cumbersome FDA regulation would ultimately be. Therefore, one can best assess the impact of the 1962 amendments by examining how innovation responded in the 1970s. The difficulty with studying the 1970s is that the US and UK regulatory systems eventually converged, so the UK was no longer obviously a valid control.[14] Therefore, investigators (Grabowski, Vernon, and Thomas, 1978; Wiggins, 1981; Wiggins, 1983, Jensen, 1987, Thomas, 1990) began quantifying FDA regulation by the amount of time it took for the FDA to review new drug applications (NDA).

Another issue that concerned economists was that, although NCE introductions fell in the 1960s, research expenditures rose. One interpretation was that the Peltzman finding underestimated the effect of FDA regulation because it focused on output rather than the productivity of research expenditures. A number of studies (Grabowski, Vernon, and Thomas, 1978; Wiggins, 1981) investigated this possibility by using NCE introductions/R&D expenditures as an outcome variable. For example, Grabowski, Vernon, and Thomas (1978) estimated that the 1962 amendments increased the average cost of each NCE by a factor of 1.86 to 2.3. In addition, Wiggins (1981, 1983) examined whether FDA regulations reduced the amount companies invested in R&D and found that delays in FDA approval due to the 1963 amendments reduced R&D expenditures in the 1970s. Holding these expenditures constant, NCE introductions fell 52%. Accounting for these reductions in R&D expenditures, NCE introductions fell a total of 135% after 1962.

[14] Because the UK still had shorter approval times, Grabowski, Vernon, and Thomas (1978) were still able to use the UK as a control, although they used approval times as a measure of FDA regulatory rigor.

While various studies have introduced other improvements to the analysis the effects of FDA regulation on innovation,[15] the most important of these is Thomas (1990), which observed that FDA regulation might have had different effects on different companies. Specifically, regulation may have had a larger effect on small companies that were unable to afford the clinical testing required by the FDA and had less experience with the FDA process than larger companies.[16] In addition, FDA regulation may have provided an indirect benefit to large companies by eliminating competition from smaller companies. As support, Thomas finds that FDA regulation did not affect NCE introductions by large firms, but did dramatically reduce NCE introduction by small firms. Moreover, due to reduced competition, sales (and market valuations) at large firms actually rose after FDA regulation.

4.2 Approval times

A second important parameter in evaluating FDA regulation is its effects on approval time. Early work by Wardell demonstrated the US drug development times grew versus the UK after the 1962 amendments (Wardell, 1973). This gap became known as the "drug lag." Thomas (1990) showed that the lag grew fastest in the 1960s, but still grew in the 1970s, despite the fact that formally the UK and US regulatory systems had converged by 1973. For example, the lag between the US and UK grew from 5 months in 1960 to 70 months in 1970, and then to 75 months by 1980.

The remaining papers that examine approval times fall into two categories. One examines heterogeneity in approval times for different drugs and the other examines the role of PDUFA in

[15] For example, May, Wardell, and Lasagna (1983) examined the number of NCEs that reached the stage of clinical testing. Cullen (1983) used companies' ratings of different countries' regulatory systems so countries other than UK might be used as controls.

[16] Carpenter, et al. (2008) provides another form of disparate impact from FDA regulation. That paper shows that the FDA takes longer to approve later drugs, giving early entrants a regulatory advantage. They find that a standard deviation increase in the log order of entry increases FDA approval time by 3.6 months. This gradient was increased by the 1962 amendments, but unaffected by PDUFA.

lowering approval times. One criticism of the early literature on drug lag was that it may overestimate the cost of FDA delay because the delay might only affect less valuable drugs. Of the studies that examine this issue, the best is Dranove and Meltzer (1994), which shows that drug approval times are lower for more important drugs, where importance is measured by FDA ranking of a drug's novelty, its commercial value once approved, its citations in the academic literature and in subsequent patents.[17]

In 1992 Congress took note of the drug lag and passed PDUFA, which imposed deadlines on the FDA's review of NDAs and provided the FDA with more resources – from user fees imposed on NDA applicants – to evaluate NDA applications more quickly. The question academics asked was whether PDUFA actually lowered approval times and, if so, whether this was due to deadlines and/or the resources provided by Congress.[18] Carpenter et al. (2003) and Olson (2004) come out on opposite sides of this debate. The difference is that Carpenter and colleagues assigned a drug to the year that its NDA application was filed,[19] while Olson assigned it to the year its NDA was approved. Since PDUFA was a national (rather than state) law, studies have used a dummy for the period after 1992 to code the treatment variable. Thus year of assignment is critical to one's findings. Olson's filings are confirmed and extended by Berndt et al. (2005a), which shows that PDUFA I (1992-2006) reduced the approval times by 7.6% annually while PDUFA II only reduced approval times by 3.6% annually. That paper also shows that, whereas PDUFA I had no effect on clinical development times, PDUFA II did lower these times by 4.5%. This is not surprising as one of the goals of subsequent versions of PDUFA was to streamline the regulatory process between the IND application and the NDA application (Hutt, Merrill, and Grossman (2007)).

[17] Another important insight in the Dranove and Meltzer study is that FDA regulation might affect not only approval times but the amount of time required for drug development. The higher the FDA standard, the more time companies have to spend investigating a drug to see if or prove it meets the higher standard. Therefore, Dranove and Melzer look at the total time from patent filing to approval for more and less important drugs.

[18] Hutt, Merrill, and Grossman (2007) report, however, that Congress reduced its funding for the FDA as user fees grew so that total funding did not grow as fast as user fees.

[19] The deadline clearly had some effect. Carpenter, Zucker, and Avorn (2008) show that PDUFA caused the FDA to make many more judgments on drugs in the two months before statutory deadlines.

4.3 Withdrawals

Early work on how FDA regulations affect the rate or time at which drugs were withdrawn from the market focused on comparing the US to the UK. They implicitly used the 1962 amendments as the treatment variable. Bakke, Wardell, and Lasagna (1984) looked at withdrawals from 1963 to 1983 and found no difference between the two countries. But this can largely be explained by the small number of withdrawals in each country and thus low power to detect any differences in withdrawal rates. Bakke et al. (1995) revisited the question with data from 1974-1993 and found a larger difference between the US and UK. As predicted the US, which had relatively strict regulation (at least as measured by approval times) had both fewer drug withdrawals (10 vs. 20 in the UK) as well as lower withdrawal rate (2% vs. 3% in the UK).

More recent work on withdrawal rates has focused on approval times as a measure for FDA regulatory intensity. Some relatively simple papers by the GAO (2002), CDER (2004) and Berndt et al. (2005b) compared the probability a drug was withdrawn during the period prior to PDUFA to the period after the statute's adoption. They uniformly found somewhat lower, but not significantly lower, withdrawal rates prior to PDUFA.

Carpenter, Zucker, and Avorn (2008a) used a more sophisticated approach to identify the effect of PDUFA. Instead of conducting a before-after PDUFA comparison, that study demonstrated that PDUFA caused the agency to compress the timing of decisions on drugs to the two months just before PDUFA deadlines (months 11 and 12 for standard review drugs, and months 9 and 10 for priority review drugs). The study then compared drugs approved close to the deadline to drugs approved well before or after deadlines. They found that drugs approved near deadlines had higher odds of being withdrawn (odds ratio = 5.5). Moreover, these drugs also had higher odds of having a black box warning (OR = 4.4) and of being voluntarily withdrawn by drug companies (OR = 3.3). Of course, these estimates only show

that earlier deadlines increase withdrawal rates. They must be divided by the change in approval times implied by the early deadline in order to generate a regulatory dose-response curve. In effect, they need to compare the timing of decisions (and withdrawals) during PDUFA to the timing of decisions (and withdrawals) prior to PDUFA.

4.4 Development costs

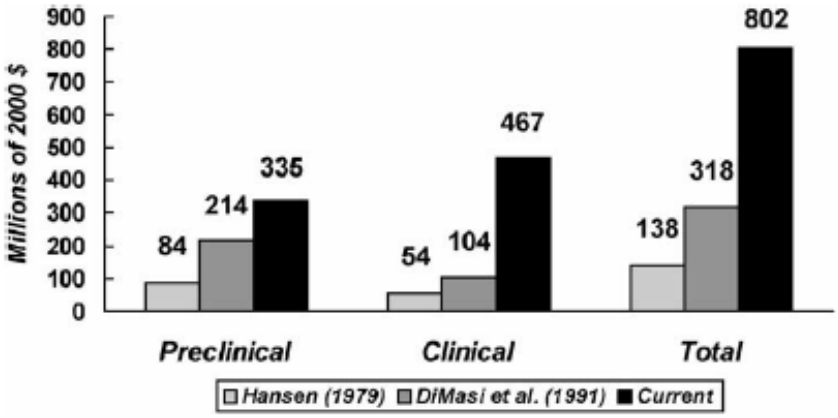
There have been a number of studies since the early 1970s that estimate the cost of drug development. These studies are spaced roughly a decade apart and generally cover the period between studies. DiMasi and Grabowski (2010) review this literature in Chapter 2. In early years, these studies relied on a small sample of drugs from a single firm (Schnee, 1972; Sarett, 1974) or aggregate data (Mund 1970; Baily 1972). More recent studies have relied on drug-level data from a sample of drug companies (e.g., Hansen, 1979; DiMasi, Bryant, and Lasagna, 1991; Adams and Brantner, 2006).[20] The latter studies attempt, on the one hand, to separate the cost of preclinical testing from clinical testing and, on the other hand, the direct out-of-pocket costs of research from the opportunity cost of that research. The last component – opportunity costs – is driven largely by delay and the real cost of capital. In order to account for the fact that many drugs ultimately fail to demonstrate value in trials, or are not approved by the FDA, the studies divide total costs by the number of drugs that are approved, resulting in an estimate of the cost per approved drug rather than, say, the cost per drug ever tested.

Together, these studies paint a picture of steadily increasing drug development costs. This is illustrated in Figure 3 below (reproduced from DiMasi, Hansen, and Grabowski, 2003), which reports estimates from Hansen (1979), DiMasi, Bryant, and Lasagna (1991), and DiMasi, Hansen, and Grabowski (2003) that roughly cover the 1970s, 1980s and 1990, respectively. Total costs per approved drug have

[20] The two major sources of data are The Tufts Center on Drug Development and Pharmaprojects.

risen from $138 million in the 1970s to $802 million in the 1990s. More recent estimates suggest the costs might now be as high as $1.6 billion per drug.

Figure 3 – Drug Development Costs



An important limitation of the literature on development costs is that it only demonstrates that costs have grown. The studies do not show that FDA regulation is responsible for this growth. While the dramatic increase in development costs after the 1962 amendments and during the run up in approval times through the 1980s suggests that the FDA is responsible, the continued growth of development costs even after the decline of approval times in the 1990s raises some questions. Has drug development hit diminishing returns and is that the main driver of cost growth in recent decades? Are approval times an adequate measure of FDA regulation or does the FDA offset lower approval times with a higher standard for minimum quality or more rigorous screening of IND applications?

4.5 Consumer and producer surplus

The final parameters required to evaluate FDA policies are consumer and producer surplus effects, ultimately driving social surplus effects. As Table 4 shows, there are only three papers that have attempted to estimate these.

Table 4. Literature concerning the social surplus from FDA regulation.

Source	Data	Measure of FDA regulation	Methodology	Finding
Peltzman (1973)	Quantity and price of prescriptions of newly introduced and old drugs, by therapeutic class and year, 1960-1962, 1964-1970	1962 Amendments	Regress market share of new drugs on ratio of new and old drug prices. Surplus is 0.5*(a – p)q, where a is the y-intercept of the demand curve estimated above.	Consumer surplus for each year's NCEs was $51.9 million/year before the 1962 amendments, $9.9m per year after the amendments. Assuming 10 percent rate of return, discounted loss from amendments was $420 million per year.
Philipson et al. (2008)	Sales for all drugs, 1998-2002; PDUFA fees	PDUFA	Regress sales on age of drug to construct age-profile of sales. Producer surplus is PV of sales – user fees – variable costs, which are ¼ to ½ of sales. Social surplus calculated as different fractions of sales (before patent expiration: all sales, ½ sales, 0; after expiration: all sales) Change in surplus from PDUFA is benefit of starting sales earlier.	Additional producer surplus from PDUFA was $8-13 billion and additional total surplus from PDUFA is $13-30 billion, assuming a 9% rate of return.
Philipson, Sun, Jena, Goldman (2009)	Survival probabilities for HIV, certain cancer patients by year; annual patient expenditures on key HIV, cancer drugs	N/A	Use Murphy-Topel framework to estimate willingness to pay for improved survival. WTP minus patient expenditures is measure of consumer surplus. Producer surplus is 80% of patient expenditures (assuming marginal costs are 20% of expenditures) Examine effect of 1 year acceleration of drug entry on social surplus.	Consumer (producer) surplus from introduction of HAART in 1996 was $364 ($38) billion. Entry 1 year earlier would have increased consumer (producer) surplus by $19 ($4) billion. Consumer (producer) surplus from introduction of Rituxan in 1998 was $12 ($4) billion. Entry 1 year earlier would have increased consumer (producer) surplus $310 ($330) million. Consumer (producer) surplus from introduction of Receptin in 1999 was $149 ($12) billion.

				Entry 1 year earlier would have increased consumer (producer) surplus $8 ($1) billion.

The first paper is Peltzman (1973a), which estimates the demand curve for new drugs by regressing the market share of newly introduced drugs in a therapeutic class on the ratio of new and old drug prices in that class. Peltzman included a dummy for the 1962 amendments as a demand shifter. Peltzman uses his estimate of the demand for new drugs only to estimate the value of information provided by the FDA minus the reduction in innovation due to FDA regulation. Peltzman's static framework for valuing the information produced by the FDA (see Section II.A.2 and Figure 5.2) suggests that the pre-1962 demand curve may not identify the "true" demand for new drugs because the FDA was not yet producing information about the quality of drugs. However, Peltzman argues that before 1962 consumers learned about the true quality of drugs through experience, and thus demand during that period was still "true" demand. He estimates that the 1962 amendments reduced demand for new drugs and thus the surplus from these drugs by roughly $420 million per year through 1970. He concludes that the loss of innovation due to the 1962 amendments offset the value of any information they provided.

The second paper to examine the FDA's impact on social surplus is Philipson et al. (2008). This differs from Peltzman (1973a) in a number of respects. Instead of studying the effect of the 1962 amendments, this paper examines PDUFA and the value of reducing FDA approval times. Moreover, the paper uses a substantially different methodology to identify surplus. Instead of estimating demand curves, the paper simply uses sales data to bind the annual social surplus from all drugs on the market during 1998-2002. It then uses drugs of different ages to estimate the stream of social surplus from a

new drug over its life cycle. Finally, it uses prior estimates of how much PDUFA accelerated drug introductions to estimate the value of accelerating these streams of social surpluses. It concludes that PDUFA, by accelerating drug approvals, increased social surplus by $13-30 billion assuming a 9% cost of capital.

The last paper is Philipson, Sun, Jena, and Goldman (2009). Like Philipson et al. (2008), the focus is identifying the value of accelerated introduction of drugs. The main difference is that the paper uses the effect of new drug introductions on survival probabilities of patients (combined with a value of life-years) to estimate a willingness to pay for a drug. Subtracting the price of the drug from this willingness to pay yields the individual patient's consumer surplus. Producer surplus is estimated as 80% of sales revenue (assuming marginal costs of 20% of revenue). After estimating the stream of aggregate social welfare from three drugs (HAART for HIV patients, Rituxan for Hodgkin's lymphoma patients, and Herceptin for breast cancer patients), the authors calculate the value of accelerating this stream by 1 year. At a 9% cost of capital, the authors estimate, for example, that introducing HAART one year earlier would have increased consumer and producer surplus by $19 and $4 billion, respectively.

Section 5: Concluding remarks and avenues of future research

We reviewed some of the central aspects of the vast positive and normative literature on the role of markets and public polices in mitigating the effects of health shocks. The literature has been primarily concerned with various forms of insurance that attempt to smooth consumption across health shocks by insuring financial effects on health care spending or wealth. It has discussed the impediments to full consumption insurance and the role of the government in addressing these impediments as well as any negative effects from the lack of universal purchase of these products. This large literature has focused almost exclusively on consumption smoothing rather than smoothing of the stock of health itself, although we argue the latter may be more important for welfare. Because human capital cannot be traded, risk pooling of health shocks is infeasible beyond the existing medical care that treats them, necessitating other forms of lowering health risk. We argued that medical innovation can be interpreted as an insurance mechanism of a population's health. We explored the positive and normative implications of this population insurance view of medical R&D and stressed the ex-ante insurance value of medical innovations.

There are several avenues of future research in examining the role of medical innovation in insuring health. One is in assessing the relative value of public subsidies for medical innovation affecting smoothing in health versus health insurance reforms affecting consumption smoothing. Much of the debate and legislation concerning health reforms has been under the rationale of reducing market inefficiencies in health-induced shocks to consumption. Our analysis may suggest that given the potentially large value of smoothing health itself rather than consumption, more explicit analysis is needed on the relative value of public programs stimulating medical innovation rather than health reforms aimed at enabling consumption smoothing.

A second area concerns a more comprehensive analysis of the role of rare disease R&D that eliminates small risks with severe health effects. Public subsidies of rare disease R&D are common, such as the Orphan Drug Act in the United States. However, according to traditional analysis they are inefficient given that small markets cannot support the fixed costs in R&D as well as larger markets. Our analysis suggests that small disease R&D may be efficient when it is interpreted as an insurance mechanism for a low probability but severe event. For the same reasons that life insurance is valuable to the vast majority of people with coverage who do not die, small disease R&D is valuable for the vast majority of people who never get the disease. More generally, the value of new medical innovations for untreated individuals never using them need to be better understood.

A third area concerns the exact risk properties of medical treatments and how FDA regulations affect them. In particular, clinical trials only estimate mean effectiveness or side effects levels, and not the covariance between them. The net benefit of a treatment—the value of health it generates net of side effects and price—has very different risk properties depending on whether side effects are positively or negatively correlated with effectiveness. If a side effect only occurs when a treatment is successful, it is a more tolerable treatment than if it only occurs when the treatment is unsuccessful. But FDA policies based on mean levels do not capture this difference in value induced by the covariance of efficacy and side effects. The overall argument is that when uncertain health can be reduced by medical R&D, the full risk properties of new treatments matter.

In general, it seems plausible that, given the large value of health relative to consumption estimated by economists (see, e.g., Murphy and Topel 2006), the current preoccupation with policies aimed at consumption smoothing across disease states may have lower marginal returns than policies aimed at smoothing health itself across those same disease states.

REFERENCES

Abaluck, J., and Gruber, J. (2011), "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review* 101(4): 1180-1210.

Abbring, J.H., P.A. Chiappori, J.J. Heckman, and J. Pinquet. 2003. "Adverse Selection and Moral Hazard: Can Dynamic Data Help to Distinguish?" *Journal of the European Economic Association* 1(2-3): 512-521.

Adams, C.P., and V.V. Brantner. 2006. "Estimating the Cost of New Drug Development: Is It Really $802 Million?" *Health Affairs* 25(2): 420-428.

Akerlof, G.A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3): 488-500.

Altman, D., D.M. Cutler, and R.J. Zeckhauser. 1998. "Adverse Selection and Adverse Retention." *American Economic Review* 88(2): 122-136.

Arrow, K.J. 1963. "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review* 53(5): 942-973.

Baily, M.N. 1972. "Research and Development Costs and Returns: The U.S. Pharmaceutical Industry." *Journal of Political Economy* 80(1): 70-85.

Bakke, O.M., W.M. Wardell, and L. Lasagna. 1984. "Drug Discontinuations in the United Kingdom and the United States, 1964 to 1983: Issues of Safety." *Clinical Pharmacology & Therapeutics* 35(5): 559-567.

Bakke, O.M., M. Manocchia, F. de Abajo, K.I. Kaitin, and L. Lasagna. 1995. "Drug Safety Discontinuations in the United Kingdon, the United States, and Spain from 1974 Through 1993: A Regulatory Perspective." *Clinical Pharmacology & Therapeutics* 58: 108-117.

Bekelman, J.E., Y. Li, and C.P. Gross. 2003. “Scope and Impact of Financial Conflicts of Interest in Biomedical Research - A Systematic Review.” *Journal of the American Medical Association* 289(4): 454-65.

Berndt, E.R., A.H.B. Gottschalk, T.J. Philipson, and M.W. Strobeck. 2005a. “Assessing the Impacts of the Prescription Drug User Fee Acts (PDUFA) on the FDA Approval Process.” *Forum for Health Economics & Policy* 8: Article 2.

Berndt, Ernst R., A.H.B. Gottschalk, T.J. Philipson, and M.W. Strobeck. 2005b. “Industry Funding of the FDA: Effects of PDUFA on Approval Times and Withdrawal Rates.” *Nature Reviews Drug Discovery* 4(7): 545-554.

Cardon, J.H., and I. Hendel. 2001. "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey." *RAND Journal of Economics* 32(3): 408-27.

Carpenter, D., M. Chernew, D.G. Smith. and A.M. Fendrick. 2003. “Approval Times for New Drugs: Does the Source of Funding for FDA Staff Matter?” *Health Affairs* hlthaff.w3.618. Available at http://content.healthaffairs.org/cgi/content/abstract/hlthaff.w3.618v1.

Carpenter, D., E.J. Zucker, and J. Avorn. 2008a. “Drug-Review Deadlines and Safety Problems.” *New England Journal of Medicine* 358(13): 1354-61.

Carpenter, D., E.J. Zucker, and J. Avorn.. 2008b. “Drug Review Deadlines and Safety Problems: Author’s Reply.” *New England Journal of Medicine* 359(1): 96-98.

Cawley, J. and T.J. Philipson. 1999. “An Empirical Examination of Information Barriers to Trade in Insurance.” *American Economic Review* 89(4): 827-846.

Center for Drug Evaluation and Research. 2004. “2003 Report to the Nation: Improving Public Health through Human Drugs.” US Dept. of Health and Human Services, Food and Drug Administration.

Chandra, A., J. Gruber, and R. McKnight. 2010. "Patient Cost-Sharing and Hospitalization Offsets in the Elderly." *American Economic Review*, 100(1): 193–213.

Chernew, M.E., and Fendrick, A.M. (2008), "Value and Increased Cost Sharing in the American Health Care System," *Health Services Research* 43(2): 251-7.

Chiappori, P.-A., and B. Salanié. 2000. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy* 108(1): 56-78.

Cohen, A. 2005. "Asymmetric Information and Learning in the Automobile Insurance Market." *Review of Economics and Statistics* 87(2): 197-207.

Cohen, A. and P. Siegelman. 2010. "Testing for Adverse Selection in Insurance Markets." *Journal of Risk and Insurance* 77(1): 39-84.

CBO. Jan 2008. Technological Change and the Growth of Health Care Spending.

Connolly, M.P., G. Griesinger, W. Ledger, and M.J. Postma. 2009. "The Impact of Introducing Patient Co-payments in Germany on the Use of IVF and ICSI: A Price-elasticity of Demand Assessment." *Human Reproduction* 24(11): 2796-2800.

Cook, P.J., and D.A. Graham. 1976. "The Demand for Insurance and Protection: The Case of Irreplaceable Commodities." *Quarterly Journal of Economics* 91(1): 143-156.

Cullen, R. 1983. "Pharmaceuticals Inter-Country Diffusion." *Managerial & Decision Economics* 4(2): 73-82.

Cutler, D.M., and M. McClellan. 2001. "Is Technological Change in Medicine Worth It?" *Health Affairs* 20(5): 11-29.

Cutler, D.M., and S.J. Reber. 1996. "Paying for Health Insurance: The Trade-off Between Competition and Adverse Selection." *Quarterly Journal of Economics* 113(2): 433-466.

Cutler, D.M., and E. Richardson. 1998. "The Value of Health: 1970-1990." *American Economic Review Papers & Proceedings* 88(2): 97-100.

Cutler, D., G. Long, E.R. Berndt, J. Royer, A.A. Fournier, A. Sasser, and P. Cremieux. 2007. "The Value of Antihypertensive Drugs: A Perspective on Medical Innovation." *Health Affairs* 26(1): 97-110.

Danzon, P.M., and M.V. Pauly. 2002. "Health Insurance and the Growth in Pharmaceutical Expenditures." *Journal of Law and Economics* 45: 587-613.

DiMasi, J.A., and H.G. Grabowski. 2010. [Handbook chapter.]

Dimasi, J.A., N.R. Bryant and L. Lasagna. 1991. "New Drug Development in the United-States from 1963 to 1990." *Clinical Pharmacology & Therapeutics* 50(5): 471-86.

Dimasi, J.A., R.W. Hansen, and H.G. Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *Journal of Health Economics* 22(2): 151-185.

Dionne, G. 1982. "Moral Hazard and State-dependent Utility Function." *Journal of Risk and Insurance* 49(3): 405-422.

Dionne, G., C. Gourileroux, and C. Vanasse (2001). "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment." *Journal of Political Economy* 109 (2): 444-453.

Dranove, D., and D. Meltzer. 1994. "Do Important Drugs Reach the Market Sooner?" *RAND Journal of Economics* 25(3): 402-423.

Fang, H., M. P. Keane, and D. Silverman. 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy* 116(2): 303-350.

Finkelstein, A. (2004), "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics* 119(2): 527-564.

Finkelstein, A. (2007), "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare," *Quarterly Journal of Economics* 122(1): 1-37.

Finkelstein, A. and K. McGarry. 2006. "Multiple Dimensions of Private Information: Evidence from the Long Term Care Insurance Market." *American Economic Review* 96(4): 938-58.

Finkelstein, A., and J. Poterba. 2002. "Selection Effects in the United Kingdom Individual Annuities Market." *Economic Journal* 112(476): 28-50.

Finkelstein, A., and J. Poterba. 2004. "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." *Journal of Political Economy* 112(1): 183-208.

Frank, R. (2007), "Behavior Economics and Health Economics," in P. Diamond and H. Vartiainen (eds.,), Behavior Economics and its Applications, Princeton University Press: Princeton, NJ.

Frank, R. and Lamiraud, K. (2009), "Choice, Price Competition, and Complexity in Markets for Health Insurance," *Journal of Economic Behavior and Organization* 71(2): 550-562.

Friedman, M. A., et al. 1999. The Safety of Newly Approved Medicines: Do Recent Market Removals Mean There Is A Problem?" *Journal of the American Medical Association* 281: 1728-1734.

Gaynor, M., J. B. Rebitzer, and L. J. Taylor, 2004. "Physician Incentives in Health Maintenance Organizations." *Journal of Political Economy* 112(4): 915-931.

General Accounting Office. 2002. "Effect of User Fees on Drug Approval Times, Withdrawals, and Other Agency Activities.

Gieringer, D.H. 1985. "The Safety and Efficacy of New Drug Approval." *Cato Journal*. 5(1): 177-201.

Goldman, D., G. Joyce, and Y. Zheng (2007). "Prescription Drug Cost Sharing: Associations with Medication and Medical Utilization and Spending and Health." *Journal of the American Medical Association* 298, 61-69.

Goldman, D., and T., Philipson (2007), "Integrated Insurance Design with Multiple Medical Technologies" , American Economic Review, March.

Grabowski, H.G., and J.M. Vernon. 1978. "Consumer Product Safety Regulation." *American Economic Review Papers & Proceedings* 68(2): 284-289.

Grabowski, H.G., J.M. Vernon, and L.G. Thomas. 1978. "Estimating the Effect of Regulation on Innovation: An International Comparative Analysis of the Pharmaceutical Industry." *Journal of Law and Economics* 21(1): 133-163.

Grace, M.F., R.W. Klein, and P.R. Kleindorfer (2004), "Homeowners Insurance with Bundled Catastrophe Coverage," *Journal of Risk and Insurance* 71(3): 351-379.

Gruber, J., and Poterba, J. (1994), "Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed," *Quarterly Journal of Economics* 109(3): 701-33.

Hansen, R.W. 1979. "The Pharmaceutical Development Process: Estimates of Current Development Costs and Times and the Effects of Regulatory Changes." In *Issues in Pharmaceutical Economics*, ed. R.I. Chien, 151-87. Lexington, MA: Lexington Books.

He, D. 2009. "The Life Insurance Market: Asymmetric Information Revisited." *Journal of Public Economics* 93(9-10): 1090-1097.

Helland, E., D. Lakdawalla, A. Malani, and S. Seabury. 2010. "The Effect of Tort Liability on the Market for Prescription Drugs." Working paper.

Hendel, I., and A. Lizzeri. 2003. "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance." *Quarterly Journal of Economics* 188(1): 299-327.

Himmelstein, D.U., D. Thorne, E. Warren, and S. Woolhandler. 2009. "Medical Bankruptcy in the United States, 2007: Results from a National Study." *The American Journal of Medicine* 122: 741-746.

Hult, K., and T. Philipson (2012), "Health Care Reforms and The Value of Future Public Liabilities", NBER WP.

Hutt, P.B., R.A. Merrill, and L.A. Grossman. 2007. *Food and Drug Law: Cases and Materials, Third Edition*. NY: Foundation Press.

Jensen, E.J. 1987. "Research Expenditures and the Discovery of New Drugs." *Journal of Industrial Economics* 36(1): 83-95.

Kaitin, K., P. DiCerbo, et al. 1991. "The new drug approvals of 1987, 1988, and 1989: trends in drug development." *The Journal of Clinical Pharmacology* 31(2): 116-122.

Keeler, E.B., and J.E. Rolph. 1989. "The Demand for Episodes of Treatment in the Health Insurance Experiment. " *Journal of Health Economics* 7: 337-367.

Kniesner, T.J., Viscusi, W.K., C. Woock, and J.P. Ziliak. 2012. "The Value of a Statistical Life: Evidence from Panel Data." Review of Economics and Statistics 94(1): 74-87.

Kolstad, C.D., T.S. Ulen, and G.V. Johnson. 1990. "Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?" *American Economic Review* 80(4): 888-901.

Koijen, R., T. Philipson, and H. Uhlig (2012), "Financial Health Economics", Working Paper, Department of Economics, University of Chicago.

Kondo, M., S. Hoshi, and I. Okubo. 2009. "Does Subsidy Work? Price Elasticity of Demand for Influenza Vaccination Among Elderly in Japan." *Health Policy* 91(3): 269-276.

Krueger, A.B., and I. Kuziemko (2013), "The Demand for Health Insurance Among Uninsured Americans: Results of a Survey Experiment and Implications for Policy," *Journal of Health Economics*, forthcoming

Lakdawalla, D., T. Philipson, and Y.R. Wang. 2006. "Intellectual Property and Marketing." National Bureau of Economic Research Working Paper 12577.

Lakdawalla, D., and N. Sood. 2009. "Innovation and the Welfare Effects of Public Drug Insurance." *Journal of Public Economics* 93(3-4):541-548.

Lakdawalla, D., and N.Sood. 2006. "Health Insurance as a Two-Part Pricing Contract." National Bureau of Economic Research Working Paper 12681.

Levy, H., and D., Meltzer, (2008), "The Impact of Health Insurance on Health." *Annual Review of Public Health*, Vol. 29: 399-409.

Liebman, J. and R. Zeckhauser (2008), "Simple Humans, Complex Insurance, Subtle Subsidies," NBER Working Paper #14330.

Lexchin, J., L.A. Bero; B. Djulbegovic, and O. Clark. 2003. "Pharmaceutical Industry Sponsorship and Research Outcome and Quality: Systematic Review." *British Medical Journal* 326(7400): 1167-70B.

Lichtenberg, F. 2003. "Pharmaceutical Innovation, Mortality Reduction, and Economic Growth." In *Measuring the Gains from Medical Research: An Economic Approach*, eds. K.M. Murphy and R.H. Topel,.74-109. Chicago: University of Chicago Press.

Malani, A., and T. Philipson (2012), "The Regulation of Medical Products", in *Handbook of The Economics of the Biopharmaceutical Sector*, edited by P. Danzon and S. Nicholson, Oxford university Press.

Malani, A., O. Benbom, and M. van der Laan. 2009. "Accounting for Differences Among Patients in the FDA Approval Process." Working paper.

Manning, R.L. 1994. "Changing Rules in Tort Law and the Market for Childhood Vaccines." *Journal of Law & Economics* 37(1): 247-275.

Manning, R.L. 1997. "Products Liability and Prescription Drug Prices in Canada and the United States." *Journal of Law & Economics* 40: 203-243.

Manning, W.G., J.P. Newhouse, N. Duan, E.B. Keeler, A. Leibowitz, and M.S. Marquis. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77(3): 251-77.

May, M.S., W.M. Wardell, and L. Lasagna. 1983. "New Drug Development During and After a Period of Regulatory Change: Clinical Research Activity of Major United States Pharmaceutical Firms, 1958-1979." *Clinical Pharmacology and Therapeutics* 33(6): 691-700.

McCarthy, D., and O.S. Mitchell. 2010. "International Adverse Selection in Life Insurance and Annuities," in S. Tuljapurkar, N. Ogawa, and A.H. Gauthier (eds.), *Ageing in Advanced Industrial States: Riding the Age Waves*, Vol. 3, Springer, 119-135.

McFadden, D. (2006), "Free Markets and Fettered Consumers," *American Economic Review* 96(1): 5-29.

McGuire, T.G. 2012. "Demand for Health Insurance," in *Handbook of Health Economics Vol. 2*, edited by M.V. Pauly, T.G. McGuire, and P.P. Barros. Elsevier.

McGuire, T.G., R. Nelson, and T. Spavins. 1975. "An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments: A Comment." *Journal of Political Economy* 83(3): 655-662.

Medtronic v. Lohr. 1996. 518 U.S. 470.

Muermann, A. and D. Straka. 2011. "Asymmetric Information in Automobile Insurance: New Evidence from Telematic Data," working paper.

Mulligan, C., and T.J. Philipson. 2004. "Insurance Market Participation under Symmetric Information," working paper.

Mund, V.F. 1970. "The Return on Investment of the Innovative Pharmaceutical Firm." In *The Economics of Drug Innovation*, ed. J.D. Cooper, 125-138. American University, Washington, DC.

Murphy, Kevin M. and Robert H. Topel. 2003. *Measuring the Gains from Medical Research: An Economic Approach*. Chicago: University of Chicago Press.

Murphy, K.M. and R.H. Topel. 2006. "The Value of Health and Longevity." *Journal of Political Economy* 114(5): 871-904.

Musgrave, R.A. 1959. *The Theory of Public Finance*. New York: McGraw Hill.

Nardinelli, C., M. Lanthier, and R. Temple. 2008. "Drug Review Deadlines and Safety Problems." *New England Journal of Medicine* 359(1): 95-96.

Newhouse, J. 1992. "Medical Care Costs: How Much Welfare Loss?" *Journal of Economic Perspectives* 6(3): 3-21.

Nyman, J.A. 1999. "The Value of Health Insurance: The Access Motive," Journal of Health Economics 18(2): 141-52.

Olson, M.K. 2004. "Explaining Reductions in FDA Drug Review Times: PDUFA Matters." *Health Affairs* http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.s1v1/DC1.

Oster E., I. Shoulson, K.A. Quaid, and E.R. Dorsey. 2010. "Genetic Adverse Selection: Evidence from Long-Term Care Insurance and Huntington's Disease." Journal of Public Economics 94: 1041-1050.

Pauly, M.V. 1968. "The Economics of Moral Hazard: Comment." *American Economic Review* 58(3): 531-537.

Pauly, M.V. 1990. "The Rational Nonpurchase of Long Term Care Insurance." *Journal of Political Economy* 98(1): 153-168.

Peltzman, S. 1973a. "An Evaluation of Consumer Safety Protection Legislation: the 1962 Drug Amendments." *Journal of Political Economy* 81(5): 1049-1091.

Peltzman, S. 1975. "An Evaluation of Consumer Product Legislation: The 1962 Drug Amendments: A Reply." *Journal of Political Economy* 83(3): 663-667.

Philipson, T.J., E.R. Berndt, A.H.B. Gottschalk, and E. Sun. 2008. "Cost-Benefit Analysis of the FDA: The Case of the Prescription Drug User Fee Acts." *Journal of Public Economics* 92: 1306-1325.

Philipson, T.J., and E. Sun. 2008. "Is the Food and Drug Administration Safe and Effective?" *Journal of Economic Perspectives* 22(1): 85-102.

Philipson, T.J., E. Sun, and D. Goldman. 2009. "The Effects of Product Liability Exemption in the Presence of the FDA." Working paper.

Philipson, T., Sun, E. Sun, A. Jena, and D. Goldman. 2009. "A Re-examination of the Costs of Medical R&D". Working paper.

Polinsky, A. M. and S. Shavell. 2010. "The Uneasy Case for Product Liability." *Harvard Law Review*. 123(6): 1437-92.

Richaudeau, D. 1999. "Automobile insurance contracts and risk of accident: an empirical test using French individual data." *The Geneva Papers on Risk and Insurance - Theory* 24 (1): 97-114.

Riegel v. Medtronic. 2008. 552 U.S. 312.

Rothschild, M. and J. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90(4): 629-49.

Saito, K. 2006. "Testing for Asymmetric Information in the Automobile Insurance Market Under Rate Regulation." *Journal of Risk and Insurance* 73(2): 335-356

Sarett, L.H. 1974. "FDA Regulations and Their Influence on Future Research and Development." *Research Management* 17(2): 18-20.

Schnee, J.E. 1972. "Development Costs: Determinants and Overruns." *Journal of Business* 45(3): 347-374.

Shavell, S. 1984a. "Liability for Harm versus Regulation of Safety." *Journal of Legal Studies* 13(2): 357-374.

Shavell, S. 1984b. "A Model of the Optimal Use of Liability and Safety Regulation." *RAND Journal of Economics* 15(2): 271-280.

Strombom, B.A., T.C. Buchmueller, and P.J. Feldstein (2002), "Switching Costs, Price Sensitivity and Health Plan Choice," *Journal of Health Economics* 21(1): 89-116.

Tabarrok, A.T. 2000. "Assessing the FDA via the Anomaly of Off-Label Drug Prescribing." *Independent Review* 5(1): 25-53.

Thomas, L.G. 1990. "Regulaton and Firm Size: FDA Impacts on Innovation." *RAND Journal of Economics* 21(4): 497-517.

Vernon, J.A. J.H. Golec, R. Lutter, and C. Nardinelli. 2009. An Exploratory Study of FDA New Drug Review Times, Prescription Drug User Fee Acts, and R&D Spending. *Quarterly Review of Economics and Finance* 49: 1260-1274.

Viscusi, W.K., and J.E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates from Around the World." *Journal of Risk and Uncertainty* 27(1): 5-76.

Viscusi, W.K. 2012. “What’s to Know? Puzzles in the Literature on the Value of a Statistical Life.” Journal of Economic Surveys 26(5): 763-768.

Wardell, W. 1973. “Introduction of New Therapeutic Drugs in The United States and Great Britain: An International Comparison.” *Clinical Pharmacology and Therapeutics* 14(5):773-90.

Washington Legal Foundation v. Henney, 56 F. Supp.2d 81 (D.D.C. 1999).

Weitzman, M.L. 1974. “Prices vs. Quantities.” *Review of Economic Studies* 41(4): 477-491.

Wiggins, S.N. 1981. “Product Quality Regulation and New Drug Introductions: Some New Evidence from the 1970s.” *Review of Economics and Statistics* 63(4): 615-619.

Wiggins, S.N. 1983. “The Impact of Regulation on Pharmaceutical Research Expenditures: A Dynamic Approach.” *Economic Inquiry* 21(1): 115-128.

Wittman, D. 1977. “Prior Regulation versus Post Liability: The Choice between Input and Output Monitoring.” *Journal of Legal Studies* 6(1): 193-211.

Wolfe, J.R., and J.H. Goddeeris. 1991. “Adverse Selection, Moral Hazard, and Wealth Effects in the Medigap Insurance Market.” *Journal of Health Economics* 10: 433-459.

Wyeth v. Levine. 2009. 555 U.S. ___.

Zeckhauser, R.J. 1970. "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives." *Journal of Economic Theory* 2(1): 10-26.

Zeckhauser, R.J. 1973. "Coverage for Catastrophic Illness." *Public Policy* 21(2): 149-72.