

Introduction

Technological change is often argued to be a central force behind the growth in health care spending.¹ Since so much of health care is publicly financed, both private and public institutions seeking to evaluate increases in their health care spending have been grappling with how to measure the value of new health care technologies brought about by R&D investments. A long-standing and vast health economics literature attempts to assess this value by the use of cost-effectiveness (also called cost-utility or cost-benefit) analysis.² This type of analysis, it is claimed, is critical for managing new technologies, their adoption, and their impact on long-term health care spending.

Although it is seldom explicitly stated as such in the literature, we argue that cost-effectiveness analysis, as predicted, is implicitly concerned with estimating the *consumer surplus* associated with a given technology, whether from the perspective of patients or of their health plans. In particular, many technology assessments attempt to quantify the health impacts of new technologies for patients or health plans by comparing the benefits accruing to patients with spending at observed market prices. Examples of such cost-effectiveness measures include spending per quality-adjusted life-year (QALY), commonly used by public health care buyers outside the United States, and value-of-life estimates, common in studies assessing the gains from increased health care spending. Although such estimates may not fully capture the unobservable aspects of consumer surplus incorporated in traditional demand estimates, standard cost-effectiveness assessment as performed in practice seems, nevertheless, to be based on that concept. In common cost-effectiveness practice, technologies are

2 INNOVATION AND TECHNOLOGY ADOPTION

deemed more valuable the more the benefits to the patient or health plan exceed what is spent on them.

However, when new technologies are brought to life through costly R&D, consumer surplus is a very poor guide for *dynamic* welfare. Since profits drive R&D, the degree to which producer surplus (that is, profits) captures social surplus, often necessarily at the expense of consumer surplus, becomes the central issue that determines dynamic efficiency. Put differently, the extent to which the net social value of a drug (the sum of the value to producers in the form of profits, and the value to consumers in the form of benefits net of spending) can be captured by producers determines the levels of R&D and dynamic efficiency. This, of course, is the rationale for the patent system, which substitutes producer surplus for consumer surplus in order to stimulate more efficient R&D investment. Therefore we argue that for the same reason that patents are preferred even though they lower consumer surplus after technologies are discovered, technology adoption criteria should be preferred that do not only focus on consumer surplus. Put differently, even though measured levels of cost-effectiveness would be higher without the patent system, since patients or health plans would then spend less to get the same technology, dynamic efficiency would clearly be lowered. An illustrative case of the dangers of cost-effectiveness criteria is that of vaccines, which in many cases have been estimated to be extremely cost-effective but, perhaps partly for that very reason, lack any appreciable R&D investment.³

In chapter 1 we motivate our argument by first documenting the extensive role of cost-effectiveness analysis in policy discussions around the world. We place special emphasis on the United States, where cost-effectiveness-based decisions have gained increased popularity among private payers and, to some extent, public payers as well. Following our analysis of the United States, we examine the use of cost-effectiveness-based technology adoption decisions in Europe generally and the United Kingdom in particular. The United Kingdom is particularly relevant given its recently introduced National Institute for Clinical Excellence, the first agency in any country to be granted the power to guide technology adoption on the basis of

“clinical and cost-effectiveness” for *all* new health technologies: pharmaceuticals, procedures, medical devices, and others. We also briefly consider the case of Australia, which in 1992 became the first country to require pharmaceutical companies to provide pharmacoeconomic assessments of all new drugs submitted for national coverage.

Chapter 2 presents our main theoretical arguments regarding the distinction between static and dynamic efficiency. Our central theme is that dynamically optimal technology assessment differs dramatically from static assessments, and in particular from cost-effectiveness assessments as performed in practice today. In fact, we argue that traditional cost-effectiveness measures should often be *minimized*, rather than maximized, in order to promote dynamic efficiency. The reason is that cost-effectiveness criteria implicitly are concerned with maximizing consumer surplus, which leads to too small a share of the total social surplus being appropriated by innovators making R&D investments. In addition, since little is understood about how cost-effectiveness criteria operate in a market context with traditional supply and demand, we examine how changes in cost and demand parameters may affect levels of cost-effectiveness observed in a market setting.

Having established that the ability of innovators to appropriate surplus from their innovations is central to dynamic efficiency, we investigate in chapter 3 the degree to which this took place in the case of a major breakthrough in medicine—the introduction after the late 1980s of new drugs to treat HIV infection and AIDS. HIV/AIDS is an important case to consider in and of itself, partly because it is the disease that claims perhaps the largest share of public sector R&D in the United States.⁴ Our major finding is that innovators captured only about 5 percent of the social surplus arising from the new HIV drugs introduced during this period: consumer and producer surpluses from these drugs amounted to roughly \$1.33 trillion and \$63 billion, respectively. Given the large estimated gains to consumers and the small share of surplus appropriated by producers, we investigate how this share varies under different valuation strategies. Our finding of a small producer share is robust to the extensions considered.⁵

4 INNOVATION AND TECHNOLOGY ADOPTION

Some simple back-of-the-envelope calculations put the small size of the innovators' share in perspective. For the consumer surplus, consider the 1.5 million U.S. citizens who have been infected by HIV since the start of the epidemic, some of whom died before drug therapy became available, whereas others lived until or contracted HIV after a series of new breakthrough drugs entered the market in the mid-1990s. Averaging across all these cohorts, the gain in life expectancy has been at least five years. Assuming a low value of a life-year of \$100,000, the added survival has thus been worth more than \$500,000 per individual and \$750 billion in the aggregate. This figure, of course, does not include the benefit to those individuals who will become infected with HIV in the future but will benefit from the drugs introduced to date; if we assume that current incidence rates persist in the future, including these individuals raises the total consumer value of these drugs above \$1 trillion.

For the producer surplus, consider that sales of HIV/AIDS drugs have grown from \$1 billion to \$4 billion annually since the breakthrough drugs came on the market in 1996. Assuming that these drugs continue to sell at current levels in the future, the present value of their total sales is about \$74 billion. We must then net out from this figure an approximation of the variable costs of production. We do this by estimating the markups stemming from differences in prices before and after the expiration of the patents on these drugs; on this basis, variable costs appear to be about 15 percent of revenue.

Given the apparently modest surplus appropriation by the producers of HIV/AIDS drugs, chapter 4 begins by examining whether producers of other health technologies with similar cost-effectiveness can expect to appropriate comparable amounts of social surplus. We develop a general result linking a technology's observed cost-effectiveness to the level of surplus appropriation by its producers. Our main findings are that cost-effectiveness is negatively related to observed surplus appropriation, and that the latter can be identified by information on a technology's cost-effectiveness and average markup. Under more-restrictive cost and demand assumptions, we use data on the observed cost-effectiveness of over 200 health

technologies to examine the distribution of surplus appropriation implied by these cost-effectiveness estimates. Interestingly, the median intervention considered has an estimated producer surplus appropriation of only 13 percent.

Given the inability of producers to appropriate much of the social value of their innovations, it seems natural to rely on the absence of market power as a potential explanation. However, we derive the surprising theoretical result that even though the *level* of profits, of course, rises as demand becomes more inelastic, the *share* of social surplus made up by profits may fall. In other words, market power may *reduce* appropriation rather than raise it. Thus, even when demand is highly inelastic (as in the case of life-saving technologies) and patients face high prices, monopolists may still capture only a small fraction of the social surplus. Since this share is crucial to dynamic efficiency, its paucity suggests a potentially large underinvestment in R&D from its dynamically optimal level.

Finally, we show that the small estimated share of social surplus captured by innovators turns out to be consistent with alternative, theory-based methods of calculating this share. As described earlier, the degree of market power (as measured by product markup) enjoyed by a firm *theoretically* implies its share of social surplus appropriated. Using price reductions upon patent expirations to estimate patent-protected markups, and hence the elasticity of demand for patent-protected drugs, we find existing levels of market power to be consistent with an innovator share of potential social surplus of about 10 percent.⁶ This is in the same order of magnitude as our *directly estimated* share of 5 percent. Given our theoretical finding that market power may *reduce* producer appropriation, the sizes of these shares are relevant to the high prices of HIV/AIDS drugs (presumably associated with a low elasticity of demand for these life-saving technologies), since they imply that producers capture a small share of the social surplus despite high prices.

Put together, our results suggest that if the new HIV/AIDS therapies are representative of other health technologies, the lack of appropriation of social surplus by innovators raises serious concerns about adherence to cost-effectiveness analysis. Despite the high

6 INNOVATION AND TECHNOLOGY ADOPTION

prices of many therapies such as the new HIV drugs, patients and health plans may be getting too good a deal in the short run, which hurts future patients in the long run by leading to lower rates of technological progress.