Altruism and Innovation in Health Care

Anupam B. Jena  Harvard University
Stéphane Mechoulan  Dalhousie University
Tomas J. Philipson  University of Chicago

Abstract

The joint presence of technological change and consumption externalities is central to health care industries around the world, because medical innovation drives the expansion of the health care sector and altruism seems to motivate many public subsidies. Although traditional economic analysis has proposed well-known remedies to deal with consumption externalities and inefficient technological change in isolation, it lacks clear principles for addressing them jointly. We argue that standard remedies to each of the two problems are inadequate. Focusing on U.S. health care, we provide illustrative calculations of the dynamic inefficiency in the level of research and development (R&D) spending when innovators are unable to appropriate the altruistic surplus of non-consumers. We calibrate that altruistic gains amount to about a quarter of consumer surplus in the baseline scenario and that R&D spending may be underprovided by as much as 60 percent.

1. Introduction

A major concern in the health care sector is balancing the altruistic externalities that motivate universal coverage with the technological change that such subsidies
may induce. Indeed, in the United States, existing evidence suggests that technological change is the key to the continued expansion of the health care sector (see, for example, Newhouse 1992), close to half of which is paid for by altruistically motivated subsidy programs such as Medicaid and Medicare. In this paper, we argue that the joint R&D-altruism problem is perhaps the central allocation problem in health care and crucial to understanding whether the observed growth in health care spending is efficient. Since developed nations implicitly have decided that it is intolerable to let people die or suffer when existing medical technologies can prevent it, public financing often covers such technologies. Yet such altruistic adoption and use of new technologies should also be evaluated in terms of the technological change they induce. It seems reasonable to argue that the long-run level of health care spending is far more influenced by these important dynamic altruistic issues than by many of the static incentives preoccupying much of health economics.

This general issue of balancing R&D and altruism is even more prominent in the subsector of health care made up of pharmaceuticals—the most R&D intensive of industries and also one often faced with human-rights-based access issues, particularly for poor nations. Indeed, the field of global health is often concerned with how to provide medical products and care to developing nations for diseases such as acquired immunodeficiency syndrome (AIDS), malaria, or tuberculosis. This global health issue concerns an R&D-altruism allocation problem that is in many ways similar to the domestic universal coverage issue in the United States.

Little explicit analysis exists, however, on the general principles that should govern appropriate policy in this area. A long-standing literature discusses efficient methods of correcting consumption externalities through applying subsidies and taxes that align private incentives with social ones, as first recognized by Pigou (1932). However, this classic problem assumes that there is no technological change in the good that confers the external effects. An equally long-standing literature tackles the appropriate methods of stimulating innovation, for example, the analysis of the welfare effects of intellectual property (IP) regulations.

However, this literature traditionally posits that there are no external effects in the consumption of the good for which there is technological change. Little is understood about the principles that should govern many important allocation problems that involve both technological change and external consumption effects.

Service Award 5 T32 GM07281) and from the Agency for Health Care Research and Quality (UCLA/RAND Training Grant T32 HS 000046).

1 Many other industries—such as research tools industries; industries with network, peer group, or herd effects; clean-energy industries; and industries in which production induces pollution—seem to involve similar issues of balancing externalities ex post with research and development (R&D) incentives ex ante.

2 Of course, there is a vast literature on the external effects of the R&D process itself rather than on the external consumption effects of the final good; see, for example, Jones and Williams (2000).

3 See Parry (1995) for an analysis of the optimal pollution tax when the state of technology is endogenous.
Given the importance of this R&D-altruism issue in health care, we analyze whether traditional economic solutions to the two problems in isolation are efficient. First, we discuss the impact of technological change on the efficiency of traditional remedies aimed at solving consumption externalities, such as altruism, in health care. We argue that classic Pigouvian solutions are inappropriate under technological change. In particular, if Pigouvian subsidies such as Medicaid appropriately reflect the ex post social value of health care consumption by the poor, they may still lead to underinvestment in R&D. For goods with external effects, just as for those without, ex post static efficiency is generally inconsistent with ex ante dynamic efficiency.

Second, we discuss the reverse problem of the impact that consumption externalities have on the appropriate stimulation of R&D. We find that standard remedies to induce technological change under altruism are inefficient. This is because such remedies focus only on consumer and producer surplus, not the surplus accruing to those nonconsumers affected externally. For example, rewards to innovation should be driven not only by profits or gains from those receiving subsidies such as Medicaid but also by altruistic surplus from those paying for these programs. Likewise, in the area of global health, the real economic gains from trade are often realized by rich nations—which provide aid for poorer ones—and the sellers of medical care. The surplus from this trade arises mostly from nonconsumers.

To consider the efficiency losses from standard remedies, we provide illustrative calibrations for the U.S. pharmaceutical market for human immunodeficiency virus (HIV) drugs and for U.S. health care markets more generally under the assumption that standard Pigouvian subsidies underlie public spending. The case of HIV is particularly relevant, as consumption of HIV drugs is financed mostly by altruistic Medicaid subsidies, and treatment underwent tremendous technological change in the mid-1990s. Our baseline calibrations imply that altruistic gains may be as high as a quarter of consumer surplus, on the order of $99 billion (in 2000 dollars) since the start of the HIV epidemic. For health care generally, our baseline calibrations suggest that altruistic surplus may again be nearly a quarter of consumer surplus, which implies estimates of just over $1.1 trillion annually. Given existing estimates of the relationship between R&D and profits, these levels of altruism imply a potential underinvestment of 23 percent for research into improved HIV therapies and 61 percent for R&D into the health care sector as a whole.

The paper is related to several literatures. First, it is related to the voluminous literature on the appropriate methods of treating externalities without technological change (see, for example, Laffont 1987; Tirole 1988). Second, the paper also extends the classic work on the trade-offs between direct R&D stimuli (push) and patents and prizes (pull) (see Nordhaus 1969, 1972; Wright 1983; Kremer and Glennerster 2004; Scotchmer 2006). Last, it relates to Weisbrod (1991), which discusses insurance affecting the type of technological change taking place but does not address the joint-allocation problem discussed in this paper.
2. Consumption Externalities and Research and Development

Consider an environment with a single potential innovation in the market. We assume that a product, if developed, has external consumption effects. To fix ideas, consider the static social surplus after the technology has been developed, given by

\[ W(y) = \pi(y) + s(y) + e(y), \]  

(1)

where \( \pi(y) \), \( s(y) \), and \( e(y) \) reflect profits, consumer surplus, and external effects induced by the output level \( y \). The expected dynamic surplus under R&D spending \( R \) is the expected static welfare less R&D spending:

\[ E[W] = P(R)W - R, \]  

(2)

where \( P(R) \) is the probability of discovery that is increasing in R&D spending. Actual R&D levels are determined by the profitability of the invention once it has been discovered and thus maximize expected profits \( P(R)\pi - R \).

2.1. Traditional Remedies to Correct Consumption Externalities

Now consider traditional Pigouvian remedies designed to solve the externality problem in consumption. These remedies aim to maximize static welfare by aligning private consumption motives with social ones, attaining the output \( y_w \) that maximizes \( W \). However, if profits drive R&D, those subsidies and taxes would be unlikely to induce the optimal level of R&D—the level that would maximize \( E[W] \) given the social value of the innovation as the reward, when maximized at \( W(y_w) \). For example, if prizes or awards are used as methods to generate profits and hence stimulate R&D, perfect competition ex post would not correct the consumption externality. Likewise, if patents are used to generate the profits, these profits would not incorporate the surplus gained by nonconsumers, and hence R&D would generally be inefficient. In general, a straightforward consequence of the theory of the second best is that a single instrument, such as a prize or a patent, cannot correct two sources of inefficiencies in output and R&D markets (Parry 1995).

The fact that static efficiency through Pigouvian measures is inconsistent with dynamic efficiency is analogous to the case of goods with only private consumption effects. Without externalities, it is well understood that efficient competition after an innovation has been discovered leads to zero profits and hence insufficient R&D incentives, which is of course the common rationale for patents. With externalities, this has the simple but unrecognized implication that Pigouvian consumption subsidies are typically inefficient under technological change. In general, arguing for Pigouvian solutions in the presence of technological change is tantamount to arguing for competitive markets for new inventions.

---

4 We ignore the debate over drug companies making more money from copycat drugs than from true innovations, making investment incentives too high; that is, we assume that innovation leads to substantive surplus creation.
Both incorrectly support static efficiency without regard to dynamic efficiency.\footnote{Note that the failure of Pigouvian solutions is not necessarily caused by the fact that patents are second-best methods of stimulating R&D. To illustrate, consider the case in which full-price discrimination among consumers is feasible so that in the absence of externalities, patents would induce a first-best allocation. However, even in that case, patents are never first best when there is an externality. This is because price discrimination does not allow the firm to capture surplus derived from nonconsumers. This implies that under a positive externality, the monopolist always underinvests in R&D. Conversely, when the externality is negative, the producer may overinvest in R&D.} Pigouvian corrections may make the total static surplus the highest, but dynamic welfare depends on the division of surplus or the incidence of Pigouvian corrections, that is, on how the distributional impact of corrections affects producers and consumers separately.

An illustration of this difference in static and dynamic efficiency is the temptation of governments to force R&D returns down after an important innovation has been discovered and altruism dictates full adoption. For example, many observers have argued that a major barrier to R&D investments in an AIDS vaccine is that developers realize that if they are successful, governments will mandate full distribution of their products at below-monopoly markups because it would be viewed as inhumane not to do so.\footnote{A similar example is related to the recent increase in avian flu. Roche Pharmaceuticals, maker of Tamiflu (a recommended treatment for avian flu), is facing significant pressure from several governments to allow generic distribution of its drug. While Tamiflu is still under patent, a number of Asian governments have threatened to bypass the patent and proceed with generic manufacturing if negotiated licensing fees are too high (Kanter 2005).} Such policies would perhaps be efficient ex post, as the developer would lose less from such price reductions than consumers and altruists gained ex post. However, this would, of course, not be dynamically efficient, as no vaccine would be developed in anticipation of this government response.

2.2. Traditional Remedies to Stimulate Research and Development

There is a large literature in economics that discusses the inefficiency in decisions involving R&D that occur when those who undertake the private cost of R&D, firms, do not receive the full social benefit of that investment (Arrow 1961; Tirole 1988; Scotchmer 2006). Under no externalities, the optimal prize is the present value of the social surplus and always dominates the optimal patent.\footnote{The exception is when the patent monopolists fully capture social surplus through price discrimination, in which case the optimal prize and optimal patent (infinite in length) yield the same dynamic welfare.} This is sometimes interpreted to mean that prizes dominate patents when there are no externalities, with the implicit assumption that the organizations selecting the prizes can set them correctly to represent social surplus. This is an assumption that many times may be unwarranted. Further, like patents, prizes have negative efficiency implications since they are financed by distor- tionary taxes on capital and labor (an issue that, for simplicity, we ignore for the present).

These discussions are incomplete, however—and the remedies implied thus
inefficient—when there are external effects in consumption of the product. This is because considering only consumer and producer surpluses as potential candidates to optimally drive R&D decisions leaves out the surplus of nonconsumers (e). Incorporating the surplus of nonconsumers as a carrot or stick for those conducting R&D is then necessary.

This is particularly relevant for global health issues—to induce efficient incentives for R&D in diseases present only in poor countries. Without externalities on rich nations, it seems efficient that a disproportionately low share of world R&D spending on drugs is allocated to third-world diseases even though these diseases may be more prevalent and clinically more devastating. Altruism or selfishly motivated externalities make the global health issue one of allocating resources under external effects of consumption and endogenous technological change.

In the presence of externalities, prizes tend to be more favored over patents the more positive the external effects are. Previously unrecognized, however, is that this dominance of prizes under positive external effects depends crucially on how production and distribution take place after the prize has been awarded. The implicit assumption of the method of production and distribution under a prize is that of free and unrestricted licensing of the patent after the discovery, hence generating the competitive output level. If prizes induce ex post efficiency without externalities, under external effects, prizes with free, unrestricted licensing and a competitive level of output may be an inefficient combination. In fact, patents may dominate prizes even under positive external effects.

For example, suppose that consumers are too poor to pay the variable costs of production, let alone the fixed costs of R&D. This implies that the social surplus consists of the external altruistic effects of richer countries. In this case, patents would induce monopoly power that would not confer any profits, and no R&D spending would take place for any patent length. The patent holder can at most only appropriate consumer surplus, which is zero when consumers cannot pay variable costs. Hence, under free licensing patents would dominate any positive prize. This is because the R&D would be undertaken without distribution, while under a patent, the R&D would not occur. Note that this has little to do with the second-best nature of patents: the problem with patents under altruism is that the output is not sold to those willing to pay for it, that is, the rich. Appropriate R&D incentives in the global health case need to take into account that the main group that benefits in an economic sense is the rich.

* Note that the effect of the size of the externality on patent length is ambiguous. For instance, a larger positive externality not only raises the social value of the invention but also increases the harm imposed by restricting its consumption through patents, creating two offsetting forces on the optimal patent life.
3. Calibration of Research and Development Inefficiencies Induced by Pigouvian Subsidies in U.S. Health Care

Given the theoretical importance of altruistic surplus for underinvestment in R&D, this section illustrates the potential size of this dynamic inefficiency for two cases: HIV drug subsidization specifically and the entire U.S. health care economy more generally. In each instance, we show the underinvestment in R&D that occurs if existing demand subsidies such as Medicaid are interpreted as Pigouvian corrections. Our estimates should be interpreted with several caveats in mind. First, the notion that the extent to which medical care is subsidized somehow reflects society’s altruism is a strong assumption: public subsidization may proceed from motives other than simply altruism. For example, interest groups representing producers or consumers may have great impact on the extent to which demand is subsidized. Observed levels of subsidization may be construed as an upper bound measure of any underlying altruism. Second, the nature of our illustrative calibration exercise requires us to use several pieces of information from different strands of literature. Since our calibrated estimates vary on the basis of the assumptions made and the point estimates used, we conduct a sensitivity analysis and display ranges of values. Ultimately, these calculations should be interpreted with caution from a quantitative standpoint, but we emphasize their qualitative implications.

3.1. Calibrating the External Consumption Effect

In the framework for static efficiency, suppose that for each unit sold, firms receive a per-unit subsidy \( \delta \) in addition to the price consumers pay for that unit, \( p(y) \). The static level of social surplus can then be written as

\[
W(y, \delta) = \left[ \frac{p(y)}{\delta} y - c(y) \right] + s(y) + [e(y) - \delta y],
\]

where the first term is profits, the second term consumer surplus, and the third the net altruistic surplus after paying for the subsidy. For a patent monopolist, the profit-maximizing output in the presence of the subsidy is

\[
y(\delta) = \arg \max \left[ \frac{p(y)}{\delta} y - c(y) \right].
\]

3.2. Parameterizing Altruism and Demand

We specify the external consumption effect \( e(y) \) to take the following form:

\[
e(y) = N\alpha s(y).
\]

This specification captures the public-good nature of the external consumption effect. That is, each of \( N \) individuals in a society is assumed to value a fraction, \( \alpha \), of the consumer surplus. Moreover, altruism is a public good in the sense

---

* For simplicity, we assume that subsidization is on a per-unit basis. Since health care is characterized by some beneficiaries receiving free care (for example, those on Medicaid) and others receiving no subsidization, we interpret our subsidies to be average subsidies to the entire population.
that each altruist’s consumption does not preclude that by another. The net surplus enjoyed by altruists is the external consumption effect less the subsidy:

$$N \left[ \alpha s(y) - \frac{\delta y}{N} \right].$$  (6)

Since each altruist pays only an $\frac{1}{N}$th of the subsidy, as the number of altruists increases, the cost to each of subsidizing a given level of output decreases.\(^{10}\)

We assume a constant elasticity of demand $q = (\beta/p)^\varepsilon$, where $\varepsilon > 0$ is the elasticity of demand and $\beta$ is a parameter that shifts demand outward.

3.1.2. Optimal Pigouvian Subsidy

The ex post Pigouvian subsidy is derived by maximizing the parameterized ex post welfare $W[y(\delta), \delta]$ with respect to $\delta$. Under constant returns to scale and a constant elasticity of demand, it is straightforward to show (see Appendix A) that the optimal subsidy, demand price, and supply price satisfy the following conditions:

$$\delta = \frac{c(1 + N\alpha)}{\varepsilon + N\alpha},$$
$$p_D = \frac{ce}{\varepsilon + N\alpha},$$
$$p_S = p_D + \delta = \frac{c(1 + \varepsilon + N\alpha)}{\varepsilon + N\alpha}.$$

The optimal subsidy is increasing in both the degree of altruism, $\alpha$, and the number of altruists, $N$. Note that the optimal subsidy in the presence of a monopolist is higher than that in perfect competition, as the monopolist restricts output. Finally, note that while the prices paid by consumers and received by firms are decreasing in $\alpha$ and $N$, firms’ profits rise with the degree of altruism and the number of altruists.

Under the assumption that the observed subsidy is the ex post Pigouvian solution to the problem of external consumption effects, the level of altruism is identified through the optimality condition:

$$\frac{\delta}{p_S} = \frac{1 + N\alpha}{1 + N\alpha + \varepsilon}.$$  (8)

\(^{10}\)The increase in $N$, through its effect on the subsidy, will increase output. Specifically, note that the quantity demanded by consumers depends on the price they face, which in turn depends on the subsidy. A lower per-person cost of subsidizing a given level of output will lead to an increase in the per-unit subsidy, $\delta$, and consequently in output. While this could possibly even lead to an overall increase in per-person costs $(\delta y/N)$, per-person costs will certainly increase above the level that would prevail if $N$ were to increase without any compensating changes in $\delta$ and $y$. 
Note that this condition implies that even in the absence of altruism, there is subsidization to correct the distortion induced by monopoly pricing.\textsuperscript{11} It is straightforward to show that under perfect competition, the analog optimality condition is

$$\frac{\delta^C}{p^C_s} = \frac{N\alpha}{N\alpha + \epsilon}. \quad (9)$$

Under perfect competition, altruism is necessary for subsidization.

### 3.2. Calibration for HIV/AIDS

Philipson and Jena (2006) estimate the consumer surplus, $s$, generated by the new HIV/AIDS technologies to be roughly $395$ billion since the start of the epidemic nearly 25 years ago. This figure is consistent with standard values of a statistical life-year around $100,000$ and a roughly 5-year extension in HIV life expectancy when averaged across all infected cohorts. In Appendix B, we discuss the methods used to estimate the share of the price that is subsidized ($\delta/p_s = .5$), the demand elasticity ($\epsilon = 1.25$), and the size of the nonconsumer pool ($N = 190$ million annually).\textsuperscript{12} In the most indirect parameter calibration, we use existing patent expiration data to estimate markups of brands relative to generic competition and hence the demand elasticity—allowing this elasticity to vary within a reasonable range does not, however, alter the qualitative predictions of our calibration.\textsuperscript{13} These quantities can then be used to identify $\alpha$, the fraction of the aggregate consumer surplus enjoyed by a single altruist, for either market structure, as well as the aggregate, external value to nonconsumers, $Na$. For the case of HIV, the aggregate value to nonconsumers is a quarter of the consumer surplus (that is, $Na = .25$). For individuals infected between 1980 and 2000, this amounts to roughly $99$ billion under the estimated level of consumer surplus.\textsuperscript{14} It is important to note that the magnitude of this effect is driven by the public-good nature of the externality. To see this more clearly, note that the aggregate external consumption effect of $99$ billion amortized over 20 years is

\begin{align*}
\text{Moreover, small observed shares are consistent with a negative external consumption effect. Since the subsidy is designed to induce a socially optimal output, if output is observed to be below the level that would be socially optimal in the absence of altruism, it must be because there is a negative externality.}
\end{align*}

\textsuperscript{12} As discussed in Appendix B, AIDS medications are largely subsidized by two programs, Medicaid and the AIDS Drug Assistance Program (ADAP), the latter administered through the federal Ryan White Comprehensive AIDS Resources Emergency (CARE) Program.

\textsuperscript{13} For example, demand for HIV/AIDS drugs may be more elastic because of the natural complementarities between life extension and the consumption of these drugs.

\textsuperscript{14} An alternative specification of the externality would be $e(y) = Na\gamma$, interpreted as altruists caring about the health of others rather than their welfare (as is true when $e(y) = Na\gamma(y)$). In this case, the share of the supply price that is subsidized, $\delta/p_s$, is equal to $[cy + Na\gamma(e - 1)]/[cy(e + 1) - Na\gamma]$. If variable costs are 20 percent of sales, $cy = 15$ billion; meanwhile, $d/p_s = .5$ and $\epsilon = 1.25$. Thus, the gross altruistic benefit ($Na\gamma$) is $2.5$ billion. In light of the $99$ billion predicted above, this result stresses the discrepancy between a wrong but commonly accepted measure of welfare—namely, health—and actual welfare.
simply $5 billion per year. With 190 million altruists enjoying this effect annually, the value of the externality amounts to $26 per altruist per year. With an estimated $3.25 billion spent on subsidies from 1980 to 2000 (50 percent of the $6.5 billion total HIV/AIDS drug spending), this amounts to $163 million spent annually by all altruists, or 85 cents per altruist per year. Including these costs of subsidization leads to a net external consumption benefit of roughly $25 ($26 – $.85) per altruist per year.

3.3. Calibration for the U.S. Health Care Sector

Recent estimates suggest that health care spending in the United States has been quite valuable, with consumer benefits of, on average, $4 to $5 for every dollar spent (see, for example, Cutler and McClellan 2001; Jena and Philipson 2008a). In general, these estimates vary significantly depending on the methods employed, the values of a statistical life-year used, and the health interventions considered (for example, interventions to reduce infant versus old-age mortality). On average, however, with nearly $1.98 trillion spent on health care in 2005 alone, this suggests an annual consumer surplus of between $5.92 and $7.89 trillion arising from health care consumption. Given our earlier results for HIV/AIDS, this raises the question of how altruistic surplus compares to consumer surplus for the health care sector as a whole.

We can use our framework to inform this question. First, since the overall market for health care (which includes hospital and physician services as well as drug therapies) is more competitive than that for HIV/AIDS, we begin by assuming that firms behave competitively—in this case, the share of the supply price that is publicly subsidized ($d/p) equals $N\alpha/(\epsilon + N\alpha)$. Second, we use the fact that Medicaid and the State Children’s Health Insurance Program (SCHIP) were the primary providers of subsidized health care in the United States as the empirical basis for altruistic spending in our model. In 2005, spending by both programs accounted for nearly 16 percent ($319 billion) of personal health care spending.

Several points are worth noting regarding our determination of which spending in the national health expenditure accounts is categorized as being altruistically motivated. First, we exclude Medicare since its benefits presumably reward contributions made by beneficiaries throughout their working lives rather than reflect purely altruistic motives on the part of the current workers. Including Medicare would simply raise the estimated level of altruism further. Second, because SCHIP has paralleled Medicaid expenditures on children, we include it as well, although

---

$^{15}$ Jena and Philipson (2008a) develop a methodology to link observed estimates of cost-effectiveness to surplus appropriation by producers. In their examination of over 200 health care technologies, the median ratio of gross benefits to spending was nearly 5, which is in line with published estimates that consumers obtain $4–$5 of benefits for every dollar spent.

$^{16}$ In 2003, prescription drugs (the lion’s share of spending being for on-patent formulations) amounted to 11 percent of U.S. health care spending. The vast majority of spending was on hospitals (32 percent), physicians (22 percent), and nursing, home health, and other professional services (19 percent).
2005 expenditures by SCHIP totaled only $8.4 billion, compared to $309 billion by Medicaid.\textsuperscript{17} Third, we include long-term care for the elderly by Medicaid in the total Medicaid figure. In 2004, Medicaid spending on long-term care for the elderly made up nearly 42 percent of total national spending on long-term care (Kaiser Commission on Medicaid and the Uninsured 2006). In most states, Medicaid financing of long-term care is intended to assist low-income individuals and those with specific functional impairments. In addition, in 2003, 37 percent of those receiving long-term care from Medicaid were under the age of 65. Because Medicaid-financed long-term care has specific requirements for income and disability, can be administered before the age of 65, and is therefore not applicable to the entire Medicare population, we include it in the total Medicaid figures representing altruistic spending.

Given the share of national health spending accounted for by Medicaid and SCHIP, we therefore assume that the share of the supply price that is publicly subsidized ($\delta/p$) equals .16, which implies that $Na = .19e$. If $e = 1.25$, the aggregate value to nonconsumers is 24 percent of consumer surplus, which is nearly identical in magnitude to our estimate for HIV/AIDS. As a benchmark case, we consider consumer surpluses arising from total health care spending that range from $5.92 to $7.89 trillion (which imply consumer benefits of $4–$5 for every dollar spent on health care). This implies an altruistic surplus of $1.41–$1.89 trillion in 2005 alone. This also corresponds to a gross benefit to each altruist of $7,500–$9,700 annually and a net benefit (gross benefit – cost of subsidy) of $5,900–$8,100.

3.4. Sensitivity Analysis

These calibrated estimates of altruistic surplus are still, of course, subject to much qualification. For example, different estimates presented in the literature of the level of consumer surplus arising from a single dollar of spending will affect our calculations. Note, however, that while the calibrated level of altruistic surplus varies on the basis of different estimates of consumer surplus, the ratio of altruistic surplus to consumer surplus identified by our model ($Na = .19e$) depends only on estimates of the elasticity of demand. To the extent that the elasticity differs from the assumed value of 1.25, both the ratio and the level of calibrated surplus will of course be affected for any given level of consumer surplus. To evaluate how the calibrated level of altruistic surplus responds to various elasticities of demand and levels of consumer surplus, Table 1 presents estimates of consumer surplus and elasticities of demand from several studies as well as the calibrated levels of altruistic surplus that those estimates imply.

Table 1 illustrates the broad range of calibrated altruistic surpluses that are possible for the level of public health subsidization observed in the United States. Depending on the elasticity of demand and the consumer surplus arising from

\textsuperscript{17} Centers for Medicare and Medicaid Services, National Health Expenditure Data (http://www.cms.hhs.gov/nationalhealthexpenddata/).
Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Consumer Surplus ($)</th>
<th>Assumed Elasticity of Demand</th>
<th>Altruistic Surplus ($) Billions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutler and Meara (2000)</td>
<td>5</td>
<td>.25</td>
<td>470</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>2,351</td>
</tr>
<tr>
<td>Cutler and McClellan (2001)</td>
<td>3–4</td>
<td>.25</td>
<td>282–376</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>1,411–1,881</td>
</tr>
<tr>
<td>Cutler, Rosen, and Vijan (2006)</td>
<td>.15–4</td>
<td>.25</td>
<td>14–376</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>71–1,881</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>470–1,411</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>2,351–4,232</td>
</tr>
<tr>
<td>Jena and Philipson (2008a)</td>
<td>4</td>
<td>.25</td>
<td>376</td>
</tr>
<tr>
<td></td>
<td>1.25</td>
<td></td>
<td>1,881</td>
</tr>
</tbody>
</table>

Note. Consumer surplus is the surplus from $1 of health spending and is based on authors’ calculations from the studies cited. Ranges of estimates are often based on the population being considered. For example, in Cutler, Rosen, and Vijan (2006), the consumer surplus of $4 per dollar spent is based on spending to reduce infant mortality, while the consumer surplus of $.15 per dollar spent is based on health spending by individuals 65 and older. The elasticity of demand of 1.25 is based on calculations shown in Appendix A. A lower point estimate of .25 is based on estimated elasticities of demand for all health services summarized in Ringel et al. (2002). Altruistic surplus is based on the authors’ calculations.

health care spending, the associated altruistic surplus calibrated from our model may vary from $94 billion to as much as $4 trillion. For elasticities of demand that are near unity and for consumer surpluses ranging between $3 and $4 for every dollar spent, the calibrated altruistic surplus is just over $1 trillion, or a fifth of the consumer surplus arising from total health care spending in the United States. Regardless of what view one may take on the specific value for money spent on health care, the ranges of values in the literature appear consistent with a reasonable prediction that the altruistic surplus generated by such spending may be quite large.

3.5. Implications for Underinvestment in Research and Development

Given the altruistic surplus implied by our model, we present back-of-the-envelope calculations on the degree of underinvestment in HIV R&D due to nonappropriation of this surplus. To do so requires two pieces of information: the amount of R&D to date and the expected increase in R&D if the altruistic surplus were fully appropriated. For the former, Jena and Philipson (2008b) report $16 billion (discounted to 1980 and in 2000 dollars) worth of private R&D into HIV/AIDS to date. For the latter, we use the estimate from Finkelstein (2004) that a $1 increase in the expected discounted present value of market revenue from a particular vaccine induces 5–6 cents’ worth of investment in that vaccine. While the relationship between expected revenues and innovation in the market for vaccines may not be directly comparable to either the market for HIV/AIDS specifically or health care generally, to our knowledge surprisingly
little empirical evidence exists on the link between profitability and R&D in health care. An exception is the relationship between market size and pharmaceutical innovation documented by Acemoglu and Linn (2004). While these authors show that a 1 percent increase in potential market size is associated with a 4–6 percent increase in new molecular entities, they do not calculate how dollar revenues map onto dollar R&D expenditures. Because of the paucity of empirical estimates linking expected revenues to R&D, our calculations should be interpreted not as definitive estimates of underinvestment in R&D but as ballpark figures of the general levels of underinvestment.

With estimates of the altruistic surplus for HIV/AIDS around $99 billion, Finkelstein’s estimates imply an underinvestment in R&D of $5 billion. These figures suggest that fuller appropriation of nonconsumer surplus would have increased R&D by 33 percent of the R&D completed to date. Put differently, our figures suggest an underinvestment in R&D of roughly 23 percent. We can compute similar estimates for health care in general, which seems all the more relevant since the U.S. Congressional Budget Office conceded in 1998 that no one knew whether current levels of pharmaceutical R&D were optimal (Outterson 2005). In 2003, private health care R&D was nearly $35 billion. With a predicted altruistic surplus of, say, $1.1 trillion in that year alone, this implies a potential increase in R&D of $55 billion, which suggests an underinvestment in overall health R&D of nearly 61 percent. With predicted altruistic surplus ranging from $94 billion to $4 trillion, the range of underinvestment in R&D could thus vary from $4.7 billion to as much as $200 billion. While these estimates rest on several strong assumptions regarding market structure, the nature of the altruism externality, and the impact of profits on R&D, they nevertheless highlight the potential magnitude of the underinvestment involved.

4. Concluding Remarks

The joint presence of technological change and altruism is central to health care industries around the world. Although traditional economic analysis has proposed well-known remedies to deal with consumption externalities and to stimulate technological change in isolation, it has not developed general principles for addressing these issues jointly. We considered the inefficiencies induced by using standard remedies to externalities and R&D stimulation when addressed in isolation. We showed the implications of these policies for the amount of underinvestment in health care R&D. In particular, our baseline illustrative calculations suggest that the aggregate value that nonconsumers place on the consumption of HIV drugs in the United States may be as high as 25 percent of the patients’ surplus, with similar estimates for health care consumption generally. For the case of HIV/AIDS, our baseline calibrations suggest that using this surplus to efficiently stimulate investment could increase R&D by as much as 33 percent.
Our simple analysis suggests several avenues for future research. While our calibrated estimates of both altruistic surplus and underinvestment in R&D appear quite large, it is important to stress that they vary within a wide range of values. Given this uncertainty and the potential magnitude of our results, perhaps the appropriate interpretation of our calibration exercise is that these empirical results must be further refined to provide accurate estimates of altruistic surplus and underinvestment in R&D.

A second area of future research could be aimed at gaining a better understanding of the efficiency properties of existing policy proposals in the area of providing health care in poor countries, the concern of global health discussed earlier. Existing policy proposals to deal with this implicit externality problem have been ad hoc in the sense that it is not clear with respect to which allocation problems the proposed solutions are optimal. Examples include Commission on Macroeconomics and Health (2001), which advocates cost-based pricing financed by donor countries, or Lanjouw (2002), who advocates country- and disease-specific cutbacks in IP rights. There is a basic conflict between these policy proposals and an efficient provision of R&D under altruism as they reduce the benefits to innovators (Philipson, Jena, and Mechoulan 2011). The rewards to innovation should be increased rather than decreased to reflect the value to altruistic nonconsumers.

Indeed, the provision of AIDS drugs in poor countries mimics the problem of providing drugs for rare diseases, as well as against agents of bioterror, in the United States, and it seems that international lessons can be learned from this domestic experience. With the purpose of stimulating R&D in disease classes too rare to generate R&D, the U.S. Orphan Drug Act of 1983 (Pub. L. No. 97-414, 96 Stat. 2049) reduced the cost and raised the benefit of R&D for such rare diseases. If a society cares or wants to provide insurance for those who are unlucky enough to catch uncommon diseases, the social surplus will, in addition to consumer surplus, contain nonconsumer benefits. The Orphan Drug Act may be interpreted as encouraging R&D to reflect altruism, as opposed to international proposals for developing-world diseases that discourage R&D in spite of such altruism. The enormous growth in the development of drugs for rare diseases generated by the Orphan Drug Act may contain important lessons for the appropriate international policy.

Some proposals even demand that shareholders of innovative firms not only fund R&D to discover new treatments but by reducing prices also cover the bill to satisfy the altruistic desires of the tax base. See also Grossman and Lai (2002), who discuss the protection of intellectual property across countries.

In the United States, the BioShield legislation authorized $5.6 billion over 10 years for the government to purchase vaccines and drugs to fight anthrax, smallpox, and other potential agents of bioterror (Project BioShield Act of 2004, Pub. Law No. 108-276, 118 Stat. 835 [July 21, 2004]). For a description of the main features of the act, see U.S. Food and Drug Administration, Developing Products for Rare Diseases and Conditions (http://www.fda.gov/orphan). Also see Grabowski (2003) for a related but independent discussion.
Last, the important issue of how world R&D should be financed across countries seems to fall under the aforementioned allocation problem. Many discussions of whether the United States is carrying too large a load of financing world drug R&D centers on the fact that about half of the world’s sales are obtained in the unregulated markets of the United States, with other price-regulated markets free riding on the R&D investments this yields. The nonexclusivity induced by the free flow of innovations across countries, and the desire to free ride due to that nonexclusivity, entail a classic externality problem in consumption ex post, with the additional feature of involving technological change.

Appendix A

Mathematics

We assume constant returns to scale (constant marginal cost $c$) and constant elasticity of demand, $p(q) = \beta q^{1/\epsilon}$. The social welfare maximization is

$$\max_y W(y) = \int_0^y p(q)dq - cy + N\alpha\left[\int_0^y p(q)dq - p(y)y\right]$$

subject to $y = y(\delta)$,

where $y(\delta) = \arg \max_y \{p(y) + \delta\} y - cy$ describes the monopolist’s optimal response to a subsidy $\delta$. Note that $p(.)$ is the price paid by the consumer and $\delta$ is the per-unit subsidy received by the monopolist above and beyond the price paid by the consumer. Under our assumptions on demand and production, it is straightforward to show that the monopolist-induced demand price and output satisfy

$$p(\delta) \equiv p[y(\delta)] = \frac{(c - \delta)\epsilon}{\epsilon - 1}, \quad y(\delta) = \left[\frac{\beta(\epsilon - 1)}{(c - \delta)\epsilon}\right]^\epsilon. \tag{A2}$$

We can rewrite the maximization in equation (A1) as follows:

$$\max_{y(\delta)} W[y(\delta)] = \int_0^{y(\delta)} p(q)dq - cy(\delta)$$

$$+ N\alpha\left[\int_0^{y(\delta)} p(q)dq - p[y(\delta)]y(\delta)\right]. \tag{A3}$$

Recalling that $p(\delta) \equiv p[y(\delta)]$, the first-order condition with respect to $\delta$ is
\[(1 + N\alpha)p(\delta) \frac{dy(\delta)}{d\delta} - [c + N\alpha p(\delta)] \frac{dy(\delta)}{d\delta} = y(\delta)N\alpha \frac{dp(\delta)}{d\delta}, \quad (A4)\]

which can be simplified to
\[ [p(\delta) - c] \frac{dy(\delta)}{d\delta} = y(\delta)N\alpha \frac{dp(\delta)}{d\delta}. \quad (A5)\]

Since, by definition, \(dp(\delta)/d\delta\) can be rewritten as \(dp[y(\delta)]/d\delta\), by the chain rule we obtain
\[ \frac{dp(\delta)}{d\delta} \equiv \frac{dp[y(\delta)]}{d\delta} = \frac{dp[y(\delta)]}{dy} \cdot \frac{dy(\delta)}{d\delta}. \quad (A6)\]

Using expression (A6), we can rewrite equation (A5) as follows:
\[ p(\delta) - c = y(\delta)N\alpha \frac{dp[y(\delta)]}{dy}, \quad (A7)\]

which, under constant elasticity of demand, can be written as
\[ p(\delta) - c = -\frac{p(\delta)N\alpha}{\varepsilon}. \quad (A8)\]

Using the expression for \(p(\delta)\) in expression (A2), we can solve equation (A8) to obtain the optimal subsidy \(\delta\) as well as the demand price \(p_D\) (recall that this is equal to \(p(\cdot)\)) and supply price \(p_s\) (note that \(p_s = p_D + \delta\)).
\[ \delta = \frac{c(1 + N\alpha)}{\varepsilon + N\alpha}, \quad p_D = \frac{ce}{\varepsilon + N\alpha}, \quad p_s = \frac{c(1 + \varepsilon + N\alpha)}{\varepsilon + N\alpha}. \quad (A9)\]

Using equations (A9), we obtain an expression relating the share of total expenditure on drugs that is publicly subsidized \((\delta/p_s)\) to the level of altruism and the elasticity of demand. Specifically,
\[ \frac{\delta}{p_s} = \frac{1 + N\alpha}{1 + N\alpha + \varepsilon}. \quad (A10)\]

Finally, we can calculate the ratio of profits to social welfare as follows:
Note that the share of social surplus appropriated to producers is positive since the monopolist operates in the elastic portion of the demand curve ($\varepsilon > 1$).

**Appendix B**

**Data**

This appendix describes how the following were obtained: (1) the share of the price of HIV/AIDS drugs that is publicly subsidized, (2) the elasticity of demand, and (3) the number of nonconsumers (altruists). These, along with consumer surplus measures obtained from Philipson and Jena (2006), were used to calibrate our model.

**Consumer Surplus from HIV/AIDS Drugs**

Using the methodology developed in Becker, Philipson, and Soares (2005), Philipson and Jena (2006) estimate the value of increased survival attributable to HIV/AIDS drugs. For each cohort infected with HIV, the authors estimate the aggregate value of improved survival relative to a benchmark case in which no treatment was available. They repeat this for each new set of cases, cohort by cohort, since the start of the epidemic and aggregate up. This delivers the gross value to consumers of improved survival induced by HIV/AIDS therapies. The consumer surplus is obtained by netting out total spending, which is described below.
Financing of HIV/AIDS Drugs

The majority of public spending on HIV/AIDS drugs is administered through two sources, Medicaid and the AIDS Drug Assistance Program (ADAP). To be eligible for Medicaid, individuals must have a low income and be in one of several mandated categories. Many AIDS patients qualify for Medicaid by being recipients of supplemental security income (SSI, one of the mandated categories). These individuals are both low income and disabled (Kates and Wilson 2004).

The AIDS Drug Assistance Program began shortly after the introduction of AZT in 1987. Since 1990, ADAP has been part of the Ryan White CARE Program, the third largest federal source for care of HIV/AIDS patients. Since 1996, Congress has specifically designated funds for ADAP through the CARE program. The AIDS Drug Assistance Program is a payer of last resort for prescription medications needed by those without insurance or other means to finance drug treatment. In 2001 alone, an estimated 135,000 individuals received assistance from ADAP.

Figure B1 presents estimates of national spending on HIV/AIDS drugs broken down by public and private payers. The estimates for total spending are from IMS Health (see Lichtenberg 2006). Public spending is approximated by the sum of Medicaid and ADAP expenditures. The Medicaid estimates include both federal and state contributions and were calculated from the Medicaid State Drug Utilization Data using national drug codes for all antiretrovirals introduced since 1987. Medicaid expenditure on HIV/AIDS drugs is unavailable prior to the last quarter of 1991—this is likely because Medicaid began its prescription drug rebate program (for all drugs, not just antiretrovirals) only in 1990. Data on ADAP expenditures are unavailable prior to 1996, although it was informally covering some individuals through the Ryan White CARE Program prior to that.

Since 1995, total spending has increased from $250 million to almost $4 billion, largely because of increased spending on protease inhibitors and nucleoside reverse transcriptase inhibitors. Figure B1 also demonstrates the large share of total spending on HIV/AIDS drugs financed by public sources, nearly 50 percent from 1996 onward. On the basis of the above data, we parameterize $\delta/p_s$ to equal .5.

---

22 Eligibility for supplemental security income requires an income below 74 percent of the federal poverty line. In 2004, this amounted to an annual income of nearly $7,000.
24 Key Milestones in CMS Programs can be downloaded at Centers for Medicare and Medicaid Services, History: Overview (http://www.cms.gov/history).
Elasticity of Demand and the Number of Altruists

We use the familiar monopolist markup condition, \((p - c)/p = 1/e\), to provide an estimate of the elasticity of demand for HIV/AIDS drugs.\(^{26}\) Using estimates from the literature on the prices of generic drugs relative to their branded counterparts, we assume variable costs to be no more than 20 percent of sales (see Caves, Whinston, and Hurwitz 1991).\(^{27}\) With constant returns to scale in variable costs, marginal cost is constant and equal to variable cost. This suggests that \((p - c)/p = .8\) or, alternatively, that \(e = 1.25\).

We assume the number of altruists financing HIV drug consumption, \(N\), to equal 190 million annually. This is the average number of adults alive in the United States each year from 1980 to 2000. While this figure does not reflect the annual number of taxpayers in the United States, it does partly capture nonworking individuals in households who also benefit from the external consumption effect. Note that our choice of \(N\) will not alter the aggregate value that altruists place on consumer surplus—it simply affects our estimates of the per-altruist external consumption benefit.

\(^{26}\) Since the monopolist produces only in the elastic portion of the demand curve, \(e\) is bounded from below by unity.

\(^{27}\) We use the price of generic drugs as an upper bound of the marginal costs of production. Caves, Whinston, and Hurwitz (1991) estimate that with 20 generic competitors, the ratio of prices between generic and branded drugs is roughly 20 percent.
References


Rosen, Allison B., David M. Cutler, Douglas M. Norton, Hsou Mei Hu, and Sandeep


