Please do not cite or distributethout permission

DYNAMIC EFFICIENCY AND PRICING OF PHARMACEUTICAL INNOVATIONS

Anirban Basu[a]

November 14, 2025

A prolonged economic debate exists on how much of the social surplus generated by an innovation should be transferred to the innovator to achieve dynamic efficiency in the innovation industry. I revisit the canonical model for pricing in innovation markets, which relies on maximizing total surplus, and contrast this model with the recently proposed alternative of maximizing consumer surplus. I set up this optimization problem as a weighted average of consumer and producer surplus to derive the conditions under which the results deviate from the canonical corner solution, where the full surplus is transferred to the manufacturer during the patent period to achieve dynamic efficiency. I empirically estimate a production function for innovations in the pharmaceutical market to determine the optimal transfer rate when the dynamic objective deviates from the canonical model. I then demonstrate, through the simulation of innovation markets under different payment regimes, that when a large and a small country coexist, the small country has the luxury of deviating from the canonical model due to free riding on the investments made by the larger country.

Keywords: Dynamic efficiency, surplus appropriation, innovation, monopoly, dynamic model.

JEL Codes: I11, I13

Word Count: 5410
No. of Figures: 5

Acknowledgements:

I thank Neeraj Sood, Chris Adams, Collin Carey, two anonymous referees, and numerous seminar participants for their helpful comments on the paper. No COI is reported. All errors are mine.

[a] Basu (corresponding author): 1959 NE Pacific St #357631, Seattle WA 98195. basua@uw.edu

Please do not cite or distributethout permission

I. INTRODUCTION

How do we determine prices for healthcare innovations without a market mechanism to facilitate these decisions? This question remains a key concern in healthcare economics. A fundamental challenge in biomedical innovation is the significant capital investment required for market entry. Recent estimates show that it takes about a billion dollars on average to bring a drug to the market. Naturally, this creates a dilemma for investors who are reluctant to take such a significant risk without guaranteed market returns when innovations emerge. In the 1980s, the United States passed laws to offer certain regulatory exclusivity to pharmaceutical patent holders. Exclusivity provides specific delays and prohibitions on the approval of competitor drugs to the patent holders whose products are approved for marketing by the Food and Drug Administration (FDA) so that the patent holder can market their drug exclusively and generate a stream of rents that suffices to generate enough profits that helps to recuperate investments in research and development (R&D) and make a profit. In other words, the laws assign specific monopoly powers to the patent holder, which would otherwise be illegal in the marketplace, for a fixed period. Such an environment provides pull-incentives for private investments in R&D. The laws, however, never specify how much profit is enough or how much of the welfare generated by the innovation should be transferred to or appropriated by the manufacturer during this exclusivity period.

The notion that higher returns usually lead to more innovation is well established. For example, the literature demonstrates that larger market sizes, either through a larger indication size of current innovation (Acemoglu & Linn, 2004; Dubois et al., 2015), insurance expansion (Blume-Kohout & Sood, 2013; Bennette et al., 2019), or higher utilization (Finkelstein, 2004), lead to additional future innovations. Finkelstein (2004) further shows that not all induced

innovations necessarily enhance welfare. More recently, Garthwaite et al. (2022) found that expanding market size is not sufficient for promoting innovation; the price paid within those markets matters. Specifically, they find that expanding Medicaid markets fail to induce additional innovation due to the low reimbursement rates of pharmaceuticals in these markets. However, it should be reiterated that the basic premise for increasing investments in innovation is driven by "pull" incentives (i.e., guaranteeing future rewards if successful innovations are brought to market), rather than "push" incentives (i.e., offering higher rewards today to encourage innovations tomorrow).

Economists have asserted the need to account for dynamic efficiencies in this market to fully maximize the long-term returns of innovation (Finkelstein, 2004; Jena & Philipson, 2008; Moreno & Ray, 2016). The canonical model in this regard was laid out by Nordhaus (1969), which stated that investments into the discovery of innovation rise with the profits expected from successful discovery. This theory was further discussed in Jena and Philipson (2008) and Lakdawalla and Sood (2012), establishing the fact that when a social planner intends to maximize the total surplus from current assets and future assets that are produced by private industry through research and development, the optimal share of surplus during the patent period that should be transferred to the innovating private sector to optimize the pull incentive is 100%. It is important to note that none of these papers formally distinguishes between the patent and post-patent periods. The general economic results (Nordhaus, 1969; Jena & Philipson 2008; Lakdawalla & Sood, 2012) derive the normative conclusion that efficient innovation is realized when the innovator receives all the social surplus generated by their innovation over the lifetime of the innovation. We will assume that this lifetime refers to the patent period set by law in any

jurisdiction.[b] Goldman et al. (2011) discuss the benefits of giving certain types of drugs longer exclusivity periods, which changes the duration over which appropriable surplus is calculated.

More recent work in health economics (Woods et al., 2024) has challenged this notion by allowing the social planner to only maximize consumer surplus, instead of total surplus, in the current and future periods and focus on a "push" incentive by making future consumer surplus dependent on today's producer surplus through today's share transfer. They estimate this dependency empirically by examining historical market productivity for pharmaceuticals in relation to the quantity and quality of innovations and find the optimal transfer share to be approximately 22%.

In this paper, we build on this debate and extend the theory to show conditions under which the optimal share transfer can deviate from the corner solution of 100% even when pull incentives are offered. Specifically, after establishing the canonical model for dynamic efficiency in Section II, we illustrate a positive scenario in Section III, based on the social planner's welfare function, which is a weighted average of consumer and producer surplus. Under this positive scenario, I show how the production function for innovation dictates the optimal share for a non-corner solution. Then, in Section IV, using empirical data for the United States on pharmaceutical innovations and R&D investments by 30 large publicly traded private firms, we estimate the production function parameters to establish a profile of optimal shares conditional on this production function. The empirical estimates, when plugged into the theoretical results, provide a reconciliation of the estimates by Jena and Philipson (2008) and Woods et al. (2024), but also highlight the distinction with Woods et al.'s approach, which is based on push incentives

[b] As explained later, an innovation continues to produce societal surplus indefinitely as all future innovations generate incremental surpluses over current technological levels.

rather than traditional pull incentives in innovation markets. Finally, in Section V, using our estimates of the innovation production function, we simulate a one-country economy and a two-country economy (of different sizes) that follow alternative innovation policies to study how total surplus and consumer surplus are realized from successful innovations over time. Section VI concludes with a discussion.

II. THE CANONICAL MODEL FOR DYNAMIC EFFICIENCY

Let a risk-neutral social payer pool risks across beneficiaries in a population by providing actuarially fair insurance regarding medical innovations. Let there be a private industry that is generally more efficient in medical R&D (compared to public R&D) and is also risk-neutral, as a whole, in bringing innovations to the market. However, the industry can only sell its products to the social payer, as demand outside insurance would be negligible. The social payer is benevolent and does not act as a monopsony. For each successful innovation, the private industry is guaranteed a patient exclusivity period during which the industry can appropriate rents from the social payer. Without loss of generality, let the social payer and the industry be engaged in a repeated game over a sequence of potential innovations/medical assets brought to market, notwithstanding several failures along the way, reflected in the industry's R&D costs.

Let, S^j represent the net present value[c] of the total surplus generated in the society from the j^{th} asset[d] brought to the market during the asset's exclusivity period. The social surplus (S) produced by an innovation is represented as

$$S = \int^{y_C} B(y) - MC(y)dy, \quad (1)$$

where $B(y)$ is the marginal social benefit function, where y ranks the individual i's marginal social benefits, $B(i)$, from the highest to the lowest. y_C is the socially competitive quantity demanded during an innovation's remaining patent life, and MC(y) is the marginal cost curve.[e] The social surplus is obtained by integrating the difference between B(y) and MC(y) over the range of y, representing the number of units purchased. Whole social surplus is represented when $y = y_C$. The marginal health benefits are valued at the social marginal willingness to pay for a unit of health. Let the marginal social benefit function (B) be given as:

$$B(i) = H(i) \cdot \theta - M(i) - \widetilde{M}(i) \quad (2)$$

where H = Health Benefit, θ = the marginal willingness to pay for a unit of health, which also reflects potential health equity considerations, M = healthcare cost-offsets other than the price of innovation, $\widetilde{M}$ = non-healthcare costs offset in society. Potential elements of $\widetilde{M}$ have been widely discussed in ISPOR Special Task Force on value assessments of medical technology (Lakdawalla et al. 2018).

The canonical model for dynamic efficiency was established by Nordhaus (1969) and later studied by Jena and Philipson (2008) and Lakdawalla and Sood (2012), who posited that the private industry maximizes expected net profits for each asset j such that:

[c] Discounting is kept implicit for clarity.
[d] Or groups of assets over time (j).
[e] The notion of surplus incorporates a value function for health units through the marginal willingness to pay for a unit of health for society.

$$V = \underset{I^j}{Max}\, f(j, E(R^j), I^j) = E(R^j) - I^j = \pi^j(I^j) \cdot \beta \cdot E(\ddot{S}^j) - I^j. \quad (3)$$

We denote partial derivatives with $()'$, a subscript for the differentiator (where specified), and star superscripts for optimal quantities throughout this paper. Here, the expected value of revenue, $E(R^j)$ depends on

(1) the probability of a successful innovation being launched (denoted as π^j), where $\pi^j = \pi^j(I^j) = \Pr(S^j > 0 \mid I^j)$ is a function of R&D dollars, I^j, invested in bringing asset *j* to the market, including all the failures along the way. We assume that the production function $\pi^j(I^j)$ is concave, with $\pi' \geq 0$ and $\pi'' \leq 0$.

(2) expected social surplus generated by the new asset, accounting for the stochastic nature of evidence uncertainty at the asset launch, i.e. $E(S^j | S^j > 0) = E(\ddot{S}^j)$, and

(3) the share of expected social surplus transferred to the industry by the social planner, β^j.

Overall, one can write $E(R^j) = \pi^j \cdot \beta \cdot E(\ddot{S}^j)$. For clarity, we drop the *j* superscript henceforth, unless otherwise specified.

From the industry perspective, the first-order condition for investment defines the canonical dynamic efficiency condition (Jena and Philipson, 2008). The FOC of (3) with respect to I ,

$$(\beta \cdot \pi_I' \cdot E(\ddot{S}) - 1) = 0, \text{ or } \quad \pi_{I^*}' = \frac{1}{\beta \cdot E(\ddot{S})} \quad (4)$$

Here, the optimality condition says that the industry will invest in research and development up to the point where the marginal dollar invested would at least fetch an expected dollar in return, as represented by the $\beta^j \cdot \pi_I^j \cdot E(\ddot{S}^j)$. Note that this investment is not conditional on the

industry's past revenue but is dependent on the industry's future expected revenue. Past revenue or profits have no bearing on the industry's decision to invest in R&D, except for the generic payer environment (i.e., features of β) that is common across time.

Applying the implicit function theorem to the first-order condition in (3) with respect to I and β, we find that

$$I'^{*}_{\beta} = 1/(\varepsilon \cdot \beta)\,, \qquad (5)$$

where $\varepsilon = -\pi''/\pi'$, the negative of the ratio of the second derivative to the first derivative of the probability of regulatory success with respect to investments at its optimal level. ε reflects the rate of decline in the marginal product of investment. For a concave production function, ε is positive. Together, (5) implies three characteristics about how the optimal investments by the industry respond to changes in β.

- Optimal level of total R&D investments increases with β.
- The marginal change in optimal investments decreases with β.
- The responsiveness of the marginal change in investments to β is larger when ε is smaller.

The first feature sets up the canonical result for optimal β from a social planner's point of view, who wants to maximize total social surplus. $\beta^{*} \to 1$, to ensure that the highest I^{*} investments by the private industry reflect the industry's optimality condition: $\pi'^{*} \cdot E\left(\ddot{S}\right) = 1$.

The second feature indicates that there is a declining impact of the pull incentive as β increases, but optimality still reaches a corner solution as the social payer wants to maximize social surplus.

The third feature suggests that improving the pull incentive is most effective when the industry is young, when the return to scale is higher. We return to this point later in our empirical production function estimates.

III. DEVIATIONS FROM THE CANONICAL MODEL

In the canonical model, the risk-neutral social planner aims to achieve dynamic efficiency by maximizing the expected total surplus for any asset *j*. Alternatively, let's consider the case where the social planner's objective function, h(), is represented by a weighted average of the expected consumer surplus and the expected producer surplus (Peltzman 1976; Polk et al. 2014; Hopenhyan and Saeedi 2023).

$$h(j, E(S^j), \beta) = \lambda \cdot (1-\beta) \cdot E(S^j) + (1-\lambda) \cdot \beta \cdot E(S^j)$$

$$= \lambda \cdot (1-\beta) \cdot \pi^j \cdot E(\ddot{S}^j) + (1-\lambda) \cdot \beta \cdot \pi^j \cdot E(\ddot{S}^j) \quad (6)$$

Here, $0 \leq \lambda \leq 1$, is the weight put on the expected consumer surplus. When $\lambda = 0.5$, the social planner puts equal weight on both counterparts and maximizes the total expected surplus. When $\lambda = 1$, the social planner maximizes only the expected consumer surplus (similar to Woods et al. 2024).

Again, suppressing the *j* subscript, and noting that π is dependent on β through its dependence on *I*, the first-order condition with respect to β from the social planner's point of view is given as:

$$\lambda \cdot [-\pi + (1-\beta) \cdot \pi' \cdot I'] \cdot E(\ddot{S}^j) + (1-\lambda) \cdot [\pi + \beta \cdot \pi' \cdot I'] \cdot E(\ddot{S}^j) = 0$$

$$\lambda \cdot \left[-\pi \cdot E(\ddot{S}^j)\right] + (1-\lambda) \cdot \left[\pi \cdot E(\ddot{S}^j)\right]$$

$$+ \lambda \cdot [(1-\beta) \cdot \pi' \cdot I'] \cdot E(\ddot{S}^j) + (1-\lambda) \cdot [\beta \cdot \pi' \cdot I'] \cdot E(\ddot{S}^j) = 0 \qquad (7)$$

Equation (7) is easy to interpret. A marginal change in β has two effects: first, it moves a portion of the surplus (=$\pi \cdot \mathrm{E}(\ddot{S})$) from the consumers to the producers; second, it induces additional investments by the private industry, which leads to the production of additional $\pi' \cdot I' \cdot \mathrm{E}(\ddot{S})$ units of surplus, which is then divided between the consumer and producer surplus based on β. Moreover, these changes to the consumer and producer surplus are now valued based on the welfare weights of λ.

It is useful to point out that, like the canonical case, the optimal share does not depend on the specific innovation that is being brought onto the market (i.e. $\mathrm{E}(\ddot{S})$), as $E(\ddot{S}^j)$ drops out in (6). The optimal share depends on the properties of the production function of the innovation and is given as:

$$\beta^* = \frac{\lambda}{(2\lambda-1)\cdot(1+\eta\cdot\varepsilon)}, \qquad (8)$$

where $\eta = \pi/\pi'$ (See Appendix for derivations). Note that when $\lambda = 0.5$, and the social planner maximizes total expected surplus, β^*reaches a corner solution as before, indicating that the industry should receive the full share. Equation (8) produces a non-corner solution only for

$\lambda > 0.5$. When $\lambda = 1$, that is, the social planner is maximizing expected consumer surplus only (as in Woods et al. (2024)), $\beta^* = 1/(1 + \eta \cdot \varepsilon)$.

IV. EMPIRICAL ESTIMATES OF THE PHARMACEUTICAL PRODUCTION FUNCTION

In this section, I estimate a production function for pharmaceutical approvals with R&D spending as input, for the first time in this literature. Simply, I try to estimate an industry-wide average estimate of $\pi(I)$, the probability of regulatory success as a function of R&D expenditures, recognizing that the production function may vary across individual companies and across individual assets within companies. While others (Scanell et al., 2012; Fernald et al. 2024; Woulters et al. 2020) only consider efficiency from the perspective of approvals per dollar of investment, I also consider the total number of assets that were being considered as part of the investment, which led to the approvals for a few. By combining these data, I estimate a formal production function that can be used in the dynamic efficiency framework. An additional byproduct of such a production function is the estimates of efficiency, which I compare to previously published estimates.

Datasets

We identified twenty pharmaceutical companies and ten biotech companies that were publicly traded and had the largest market caps in 2024 (listed through Yahoo Finance, spanning multiple US exchanges). We extracted information from each of these companies' annual 10-K filings from 2005 through 2023. Specifically, we extracted data on their annual R&D spending.

We also collected data on their acquisition costs. However, many did not distinguish these costs between R&D and prospects, so we did not include them in our analysis. All amounts were converted to US dollar equivalent and expressed in 2023 real dollars.

Next, using the clinicaltrials.gov platform, we extracted information on all unique drug-dosage-form assets in early-phase to Phase 3 industry-sponsored trials for these specific companies, along with the start-to-end date for each trial. We let each asset enter the cumulative "risk set" for a company, starting with the earliest date of any trials. These data contributed to forming the denominator (Y_2) of regulatory success probability.

Lastly, we extracted annual data from the FDA Orange Book on unique drug-dosage-form approvals by the FDA for every company. These data contributed to forming the numerator (Y_1) of the regulatory success probability and were merged with the clinicaltrials.gov data by company and year.

Statistical Analyses

Denoting the earliest year of data for each company as year 1, we created three time-varying (annual, y) variables for each company (m):

$I_{m,y}$ = Scaled cumulative R&D Expenditures of company m in year y

= Cumulative R&D expenditures of company m in year y

- R&D expenditure of company m in year 1. (9)

$Y_{1m,y}$ = Scaled cumulative # of approvals of company m in year y

= Cumulative # of approvals of company m in year y

- # of approvals of company *m* in year 1. (10)

$Y_{2m,y}$ = Scaled cumulative # of assets of company *m* in year *y*

= Cumulative # of assets of company *m* in year *y*

- # of assets of company *m* in year 1

- Scaled cumulative # of patents of company *m* in year *y-1*, and (11)

$\pi_{m,y}$ = $Y_{1m,y}$ / $Y_{2m,y}$ = Probability of regulatory success (12)

By definition, therefore, $I_{m,1} = Y_{1m,1} = Y_{2m,1} = 0$. $\pi_{m,1}$ remains undefined. Therefore, in year 1, we discounted the effects of all previous R&D expenditures, any assets developed previously, and any approvals coming out that year. Surely, R&D expenditures incurred before year 1 may continue to affect approvals in future years. However, these influences were controlled using the fixed-effects estimator to estimate the production function,

$E(\pi_{m,y+1}|\,.) = \alpha_{0m} + \alpha_1 \cdot \ln(I_{m,y})$, where (13)

$\ln(I_{m,y})$ = 0 when $I_{m,y}$=0, and

$\alpha_{0m} = \alpha_{00m} + \alpha_{01} \cdot Ind(I_{m,y}=0) + \alpha_{02} \cdot Y_{2m,y+1}$

Here, α_{0m} include a company-specific fixed effect (i.e., α_{00m}, which absorbs the company-specific calendar year at year 1, R&D expenditures incurred before year 1, and every other company-specific durable characteristics), an indicator for R&D investments greater than zero from year 1 (this applies in year 1 only), and the scaled number of assets in year *y+1* (i.e. the denominator of the outcome being modeled).

Under our primary specification of the log-concave production function, ignoring the fixed stock represented by the intercept, we have $\pi_{m,y+1} \sim \alpha_1 \cdot \ln(I_{m,y})$, such that $\eta = \pi/\pi_I = I_{m,y} \cdot \ln(I_{m,y})$, and $\varepsilon = -\pi_{II}/\pi_I = 1/I_{m,y}$. Therefore, η and ε are not dependent on regression coefficients. For robustness, we also explored an alternative production function where $\ln(I_{m,y})$ in (12) was replaced by $(I_{m,y})^2$. Model selection was based on AIC and BIC criteria.

We derive empirical estimates of $E(1/(1+\eta \cdot \varepsilon))$ from our production function. Under the log-concave model, $E(1/(1+\eta \cdot \varepsilon)) = E(\ln(I_{m,y})/(1+\ln(I_{m,y})))$.

We also study estimates of α_1. Given the number of assets at any time for any of the innovators, we predict the expected number of approvals at any given level of R&D investment based on the estimated production function. We then summarize these expected approvals as a function of investments using linear regression. We compare the estimated slope from this regression to estimates in the literature on the returns to a billion dollars of R&D investments. The inverse of this slope also provides an estimate of the average R&D investment required to bring an innovation to the market.

Results

Of the thirty companies for which data were extracted, nine were dropped as they did not have at least ten years of follow-up data, one additional company was dropped as we could not separate out their pharmaceutical R&D expenditures from non-pharmaceutical expenditures, another company was dropped due to insufficient trial data, and finally, two additional companies were dropped due to insufficient approvals data. The final analytical sample consisted of 17 companies. The trends in $I_{m,y}$ and in $\pi_{m,y}$ are shown in Figures 1(a) and 1 (b).

The model in (13) was estimated using these data and yielded a regression coefficient estimate of α_1 = 0.019 (std. error = 0.008; 95% CI: 0.002, 0.037).[f] The estimated concave production function and the underlying data are illustrated in Figure 2. The profiles η and ε over R&D investments based on the estimated log-concave production function are shown in Appendix Figure A1. $E(\ 1/(1+\eta\cdot\varepsilon))$ was estimated to be 0.26. Consequently, and following (8), Figure 3 illustrates the optimal transfer share to the innovator, depending on the relative weight that the social planner assigns to consumer surplus. The optimal share reaches the corner solution of 100% till the relative weight is 0.5 (which illustrates the Jena and Philipson scenario of maximizing total surplus) and then decreases as the social planner puts more weight on consumer surplus, culminating to an optimal share of 26% when the social planner only maximizes consumer surplus (illustrating Woods et al. scenario).

The profiles for the expected number of innovations given the estimated log-concave production function and the observed temporal trends of assets under R&D for each innovator are shown in Appendix Figure A2. A linear regression of these quantities over R&D investments yields a slope coefficient of 0.25 (95% CI: 0.05, 0.45). This indicates that the industry is generating 0.25 innovations per billion dollars of R&D investment. This estimate aligns squarely with previously published industry-wide estimates in Scanell et al. (2012) and Fernald et al. (2024). This also implies that the average R&D costs of bringing an innovation to the market are about $4 billion (95% CI: $2 billion, $22 billion). This estimate also aligns well with other estimates in the literature, including Dimasi et al. (2016) and Wouters et al. (2020).

[f] The model with $\ln(I_{m,y})$ fit the data better than the model with $(I_{m,y})^2$. The AIC and BIC criteria for the former were -1026.5 and -1016.1 versus -945.9 and -938.9 for the later.

Please do not cite or distributethout permission

V. IMPLICATIONS FOR INNOVATIONS AND WELFARE

Based on the estimates of the production function and the theoretical results for the pull incentive and private R&D investments, we simulate four scenarios to compare their welfare implications.

For each of these scenarios (described below), the following common assumptions are made:

1. The industry has an equilibrium pipeline of about 100 assets every year.
2. Each of the 100 assets' annual total surplus is drawn from a distribution ~ Gamma(shape = 0.6, scale = 2) (with a mean = \$1.2B, median = 0.6B, and variance = \$2.88B) for a large country (roughly based on sales data from Wouters et al. 2024). These values are curtailed at \$0.3B at the bottom. Let the commensurate annual total surplus value be half of that for a smaller country. The difference in valuation between these countries stems from their population size and GDP per capita.
3. Based on the optimality condition in (4) and the production function estimated, the asset-specific R&D is given as: $I^* = \alpha_1 \cdot \beta^* \cdot E(\ddot{S})$. We estimate this using a stochastic draw for α_1~Normal (0.019, std. error = 0.008), different values of β^* , and $E(\ddot{S})$ calculated as the sum of the annual total surpluses of each asset over a 10-year period (mimicking patent life), discounted at 3%.
4. Next, we calculate the industry-wide probability of regulatory success based on the total industry-wide R&D investments (summed over all potential assets in that year) using our

estimated industry-wide average production function. We then use this estimated probability to stochastically select which of the 100 assets enter the market each year.

5. If an asset does enter the market (after regulatory success), then it maintains monopoly power for 10 years, after which, it is assumed that the consumer surplus equals the total surplus.
6. During the 10 years of monopoly power, the consumer surplus and producer surplus are split based on the estimate of β^* used by a social planner.

With these common assumptions, we study four different scenarios reflecting different innovation policies.

A. Only one large country exists, and it maximizes total surplus, i.e. $\beta^* = 1$
B. Only one large country exists, and it maximizes consumer surplus, i.e. $\beta^* = 0.26$
C. A large and a small country coexist. The large country maximizes total surplus, i.e. $\beta^* = 1$. But the small country maximizes consumer surplus, i.e. $\beta^* = 0.26$
D. A large and a small country coexist. Both countries maximize total surplus, i.e. $\beta^* = 1$.

The expected welfare growth under these four scenarios and the common assumptions are studied over a period of 100 years, to inform a risk-neutral social planner in each country.

Results

Under Scenario A, about 45.6 (sd = 0.56) innovations come to the market every year for the sole large country, compared to only 32.7 (sd = 0.58) innovations per year under Scenario B. Figure 4a shows the trend in the net present value (NPV) of the annual total surplus generated under Scenarios A and B. As expected, the total surplus under Scenario A is much higher than B.

The NPV of the difference in total surpluses over a 100-year period amounts to $15.5 trillion. Figure 4b illustrates the trends in the annual consumer surplus. Obviously, given the assumed remaining patent duration of 10 years for all assets, consumer surplus is zero during the first ten years under Scenario A. However, starting from the 11th year, the consumer surplus becomes positive as innovations that emerged in the last 10 years begin to go off patent. We estimate that the NPV of the annual consumer surplus under Scenario A overtakes that under Scenario B around the 30th year mark. The NPV of the difference in consumer surpluses between the two scenarios over a 100-year period amounts to $2.1 trillion. Therefore, maximizing consumer surplus during the patent period does not translate to maximizing consumer surplus generally over a longer time horizon.

Figure 5 illustrates the results for scenarios C and D, where two countries (one large and the other small) coexist. Compared to Scenario C, the NPVs of both the annual total surplus and the annual consumer surplus are higher for the larger country under Scenario D (Figures 5a and 5b). The smaller country paying the manufacturer the whole surplus during the patent period adds about three additional innovations to the market each year (46 vs 49), and increases the larger country's total and consumer surplus over a 100-year period by $2.9 trillion and $2.0 trillion, respectively. Figures 5c and 5d illustrate the consequences for the smaller country. Scenario D does produce a slightly higher total surplus for the smaller country (100-year difference of + $1.4 trillion), but not a higher consumer surplus (100-year difference of - $5.2 trillion or 20% reduction). This result highlights the contrast between Scenarios A and B versus Scenarios C and D. The smaller country, when coexisting with a larger country, does not have an incentive to pay the full surplus to the innovators during the patent period if the goal of the country is to maximize consumer surplus. The presence of a larger country that drives

innovation is sufficient for the smaller country to reap the benefits of innovation without having to invest at the same level as the larger country in innovation costs. These results are robust to a wide range of alternate parameter estimates, as illustrated in Appendix Table A1.

VI. DISCUSSIONS AND CONCLUSION

There is an ongoing debate about the level of reimbursement necessary to achieve dynamic efficiency in the innovation markets. In the US, pharmaceutical innovators have been able to price the innovations they bring to the market freely. More recently, however, these prices have been scrutinized, and the Inflation Reduction Act has allowed negotiations for certain drugs in the market that are still under patent. The larger debate about pricing involves health economists in different countries pursuing different objectives for social payers and recommending different "appropriation rates" of the total surplus generated by innovations during their patent period for the innovators. In this paper, I lay out this debate by extending the canonical model of dynamic efficiency, as outlined by Nordhaus (1969) and by Jena and Philipson (2008), to incorporate a broader class of welfare functions, thereby identifying the conditions necessary to deviate from the optimal appropriation rates in the canonical model. I demonstrate that the optimal appropriation rate deviates from the canonical result of 100% only when a social payer places more weight on consumer surplus than producer surplus.

The optimal appropriation rate, then, depends on parameters of the production function for innovation, which I estimate using empirical data. Plugging the production function estimates into the theoretical results reveals that the optimal appropriation rate, when the social payer maximizes consumer surplus, is approximately 26%.

Although this estimate is close to the Woods et al. estimate of 22%, there are several differences between our results. Woods et al. derive their estimate based on push incentives, which assume that current profits inform R&D investments for future products. However, most economic models for innovation reject such push incentives. Such incentives cannot explain why venture capitalists enter the biotechnology markets.

Why a society chooses to maximize consumer surplus can be driven by various political and social factors. But would a lower-than-100% appropriation rate for industry maximize not only total surplus but also consumer surplus in the long run? To understand this question, I set up a one-country economy and a two-country economy (of different sizes) that follow alternative innovation policies, studying how total surplus and consumer surplus are realized from successful innovations over time. The main result from these simulations demonstrates that in the presence of a larger country that drives innovation by maximizing total surplus, a smaller country that aims to maximize consumer surplus would offer less than 100% appropriation rate to industry as it reaps the benefits of innovations without having to invest at the same level as the larger country in innovation costs. This could explain some of the price differences in the US and other countries (Danzon and Furukawa, 2008). This also illustrates the basis of free riding in the innovation market (Kyle 2025). One solution to this conundrum would be to internationalize prices among certain developed countries, as recently proposed by Ho and Pakes (2025). However, differences in GDP and the per capita surplus levels across countries must be considered in such policies. To this extent, formal calculations of expected surpluses would remain essential to enforce any such policies.

An important regulation relevant to this discussion is the imposition of reference pricing by the large country. For example, in 2025, the US administration issued executive orders for

instituting “Most Favored Nation” (MFN) pricing, especially for the Medicaid population. A key goal for the USMFN is to stop other (rich) OECD countries from freeloading on US investments in innovation. By referencing the low prices available in other jurisdictions, the referrer, with a market as large as the US, will almost guarantee upward pressure on supply prices in those other jurisdictions. Consequently, manufacturers would rather not launch a product in those other markets if forced by the Health Technology Assessment (HTA) committees in those countries to give a steep discount that would erode the manufacturer's profits in the US through MFN pricing. This can be empirically tested over the next few years. However, this raises a core gamble that MFN will invoke for innovation overall.

Suppose the other countries' HTA bodies succumb to the public pressure for access to drugs in a timely fashion. In that case, launch prices in those countries will inevitably rise as they prepare to offer a higher portion of the total surplus to the innovators. This rise in foreign drug prices will increase manufacturer profits through global sales, thereby signaling a pull incentive for additional innovations in the market. In contrast, if the HTA bodies in other countries do not succumb to pressure and adhere to their strict evaluation criteria for drug value, manufacturers may delay or even decide not to launch a new drug, fearing being referenced by the US. Such missing markets would lead to a loss of revenue for the manufacturers and a dampening of incentives for global innovation, including the US. Either scenario will not significantly impact US prices or the absolute levels of industry profits from the US market, as also asserted by other researchers (Rand and Kesselheim, 2021; Dubois et al., 2022). Assuming that global prices can hardly be kept hidden given manufacturers' public annual securities filings, the US MFN pricing will either reference much higher prices in other countries or not have a comparable price to reference, thereby sustaining current pricing practices in the US.

There are some limitations to our analysis. We aimed to determine an average production function at the industry level, encompassing all assets invested in the companies. Given that such a production function estimate did not exist in the literature, this was a first attempt to establish one. However, a more granular production function for different kinds of assets, e.g., small molecules versus cell and gene therapies, would also be useful. The granularity needed for such analyses was not available in our data. Many companies did not report R&D expenditure by specific assets. And even if they did, the asset names do not translate to the numeric drug names used in the clinical trials registry. We hope that better data in the future will help establish granular production functions.

I hope these results and discussions shed new light and continue to shape the debate about the level of reimbursement necessary to achieve dynamic efficiency in innovation markets.

APPENDIX

Derivations for optimal appropriation rate.

Following (7), the first order condition of the weighted objective function in (6) with respect to β is given as :

$$\lambda \cdot \left[-\pi \cdot E(\ddot{S}^j)\right] + (1-\lambda) \cdot \left[\pi \cdot E(\ddot{S}^j)\right]$$

$$+ \lambda \cdot \left[(1-\beta) \cdot \pi' \cdot I'\right] \cdot E(\ddot{S}^j) + (1-\lambda) \cdot [\beta \cdot \pi' \cdot I'] \cdot E(\ddot{S}^j) = 0$$

Canceling out $E(\ddot{S}^j)$ from each of these terms, we have

$$-\pi \cdot \lambda + (1-\lambda) \cdot \pi + \lambda \cdot \left[(1-\beta) \cdot \pi' \cdot I'\right] + (1-\lambda) \cdot [\beta \cdot \pi' \cdot I'] = 0$$

$$\pi \cdot [1 - 2 \cdot \lambda] + \lambda \cdot \pi' \cdot I' + \beta \cdot [(1-\lambda) \cdot \pi' \cdot I' - \lambda \cdot \pi' \cdot I'] = 0$$

$$\pi \cdot [1 - 2 \cdot \lambda] + \lambda \cdot \pi' \cdot I' + \beta \cdot \pi' \cdot I' \cdot [1 - 2 \cdot \lambda] = 0$$

Replacing optimal $I'^* = 1/(\varepsilon \cdot \beta)$ from (5), we have

$$\pi \cdot [1 - 2 \cdot \lambda] + \frac{\lambda \cdot \pi'}{\varepsilon \cdot \beta} + \frac{\pi'}{\varepsilon} \cdot [1 - 2 \cdot \lambda] = 0$$

$$\frac{\lambda \cdot \pi'}{\varepsilon \cdot \beta} = [2 \cdot \lambda - 1] \cdot \left[\pi + \frac{\pi'}{\varepsilon}\right]$$

$$\beta^* = \frac{\lambda \cdot \pi'}{\varepsilon \cdot [2 \cdot \lambda - 1] \cdot \left[\pi + \frac{\pi'}{\varepsilon}\right]} = \frac{\lambda}{[2 \cdot \lambda - 1] \cdot \left[\frac{\pi \cdot \varepsilon + \pi'}{\pi'}\right]} = \frac{\lambda}{[2 \cdot \lambda - 1] \cdot [1 + \eta \cdot \varepsilon]},$$

where $\eta = \pi/\pi'$.

Figure A1: Profiles of η and ε over R&D investments based on the log-concave production function $\pi \sim 0.019 \cdot \ln(I)$.

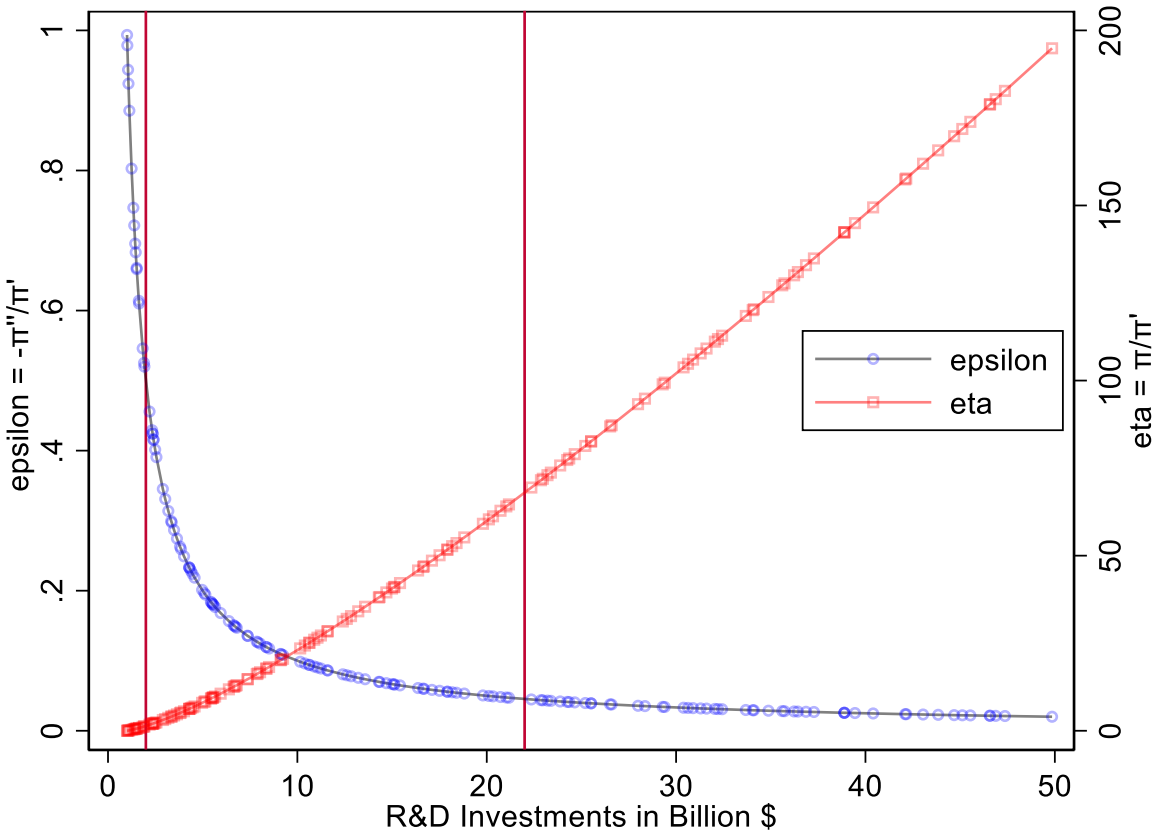


Figure A2: Linear regression (and 95% confidence band) for the Expected number of innovations given the estimated log-concave production function and the observed temporal trends of assets under R&D for each innovator, as a function of R&D investments.

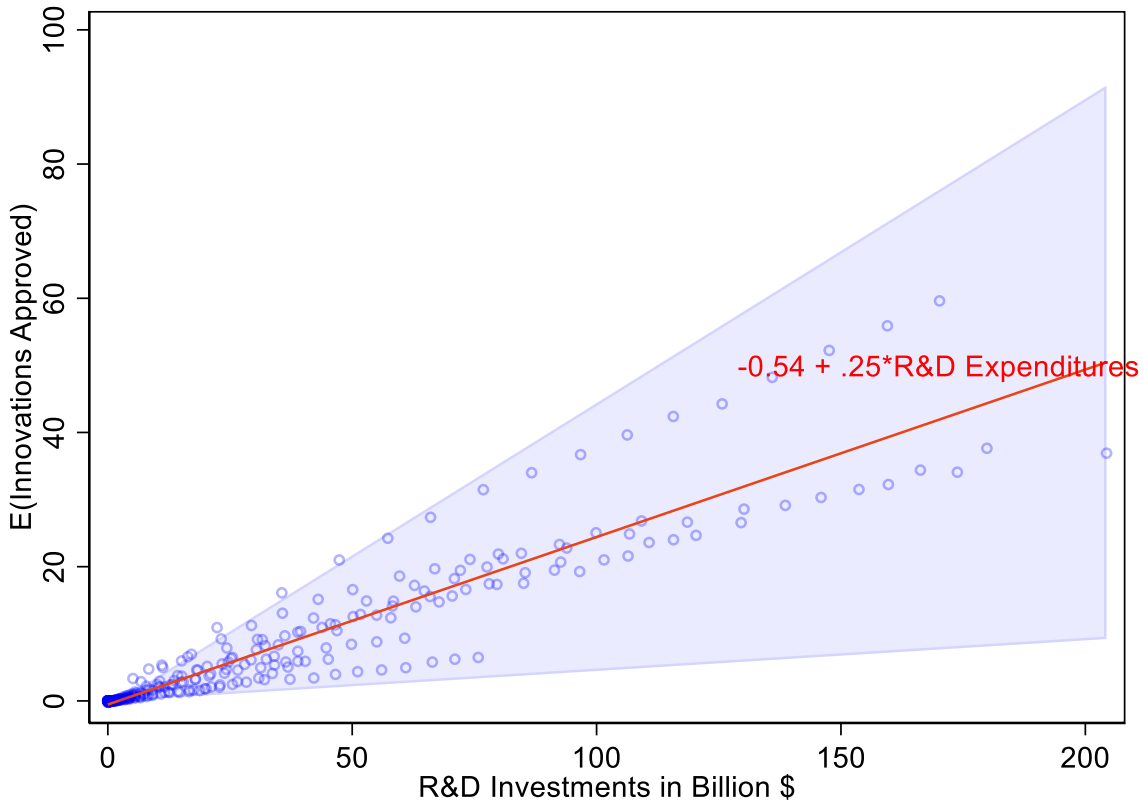


Table A1: Sensitivity analysis on the free-riding result for the small country comparing innovation policy C (A large and a small country coexist. The large country maximizes total surplus, i.e. $\beta^* = 1$. But the small country maximizes consumer surplus, i.e. $\beta^* = 0.26$) and policy D (A large and a small country coexist. Both countries maximize total surplus, i.e. $\beta^* = 1$), under alternate scenarios

Parameters	Cumulative NPV of Consumer surplus for a small country over 100 years (in $Trillions)		Difference	% change compared to Scenario C
	Under Scenario D	Under Scenario C		
Base-case Scenario+	20.4	25.7	-5.2	-20
Sensitivity to # annual potential assets				
250	8.9	11.0	-2.1	-19
1000	46.5	58.4	-11.8	-20
Sensitivity to patent length after approval				
5 years	23.8	25.6	-1.8	-7
20 years	15.2	26.2	-11.0	-42
Sensitivity to the distribution of surplus benefits for potential assets				
Uniform(0.3, 2)	18.3	22.9	-4.6	-20
Uniform(0.3, 10)	105.9	134.0	-28.1	-21
Sensitivity to the discount rate				
0%	127.8	141.5	-13.6	-10
7%	3.7	5.9	-2.3	-38

+ Base-case scenario: # Annual potential assets = 500; Patent length after approval = 10 years; Distribution of surplus benefits for potential assets = Gamma(0.6, 2); Discount rate: 3%

REFERENCES

Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3): 1049–90.

Bennette, Colleen S., Scott D. Ramsey, Zachary Helms, and Peter B. Bach. 2019. "Health Returns to Pharmaceutical Innovation in the Market for Oral Chemotherapy in Response to Insurance Coverage Expansion." *American Journal of Health Economics* 5 (3): 360–75.

Blume-Kohout, Margaret E., and Neeraj Sood. 2013. "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development." *Journal of Public Economics* 97: 327–36.

Danzon, Patricia M., and Michael F. Furukawa. 2008. "International Prices and Availability of Pharmaceuticals in 2005." *Health Affairs* 27 (1): 221–33.

DiMasi, Joseph A., Henry G. Grabowski, and Ronald W. Hansen. 2016. "Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs." *Journal of Health Economics* 47: 20–33.

Dubois, Pierre, Aditi Gandhi, and Suellen Vasserman. 2022. "Bargaining and International Reference Pricing in the Pharmaceutical Industry." NBER Working Paper 30053. Cambridge, MA: National Bureau of Economic Research.

Dubois, Pierre, Olivia de Mouzon, Fiona Scott Morton, and Paul Seabright. 2015. "Market Size and Pharmaceutical Innovation." *RAND Journal of Economics* 46 (4): 844–71.

Fernald, Kristin D. S., Philipp C. Förster, Erik Claassen, and L. H. M. van de Burgwal. 2024. "The Pharmaceutical Productivity Gap: Incremental Decline in R&D Efficiency Despite Transient Improvements." *Drug Discovery Today* 29 (11): 104160.

Finkelstein, Amy. 2004. "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry." *Quarterly Journal of Economics* 119 (2): 527–64.

Garthwaite, Craig, Rachel Sachs, and Andrew D. Stern. 2022. *Which Markets (Don't) Drive Pharmaceutical Innovation? Evidence from U.S. Medicaid Expansions.* NBER Working Paper 28755. Cambridge, MA: National Bureau of Economic Research.

Goldman, Dana P., Darius N. Lakdawalla, Jesse D. Malkin, Joel Romley, and Tomas Philipson. 2011. "The Benefits from Giving Makers of Conventional 'Small Molecule' Drugs Longer Exclusivity over Clinical Trial Data." *Health Affairs* 30 (1): 84–90.

Ho, Katherine, and Ariel Pakes. 2025. "Policy Options for the Drug Pricing Conundrum." *Proceedings of the National Academy of Sciences* 122 (9): e2418540122.

Hopenhayn, Hugo, and Mohammad Saeedi. 2023. "Optimal Information Disclosure and Market Outcomes." *Theoretical Economics* 18 (4): 1317–44.

Jena, Anupam B., and Tomas J. Philipson. 2008. "Cost-Effectiveness Analysis and Innovation." *Journal of Health Economics* 27 (5): 1224–36.

Kyle, Margaret K. 2025. "Lessons for the United States from Pharmaceutical Regulation Abroad." *Journal of Economic Perspectives* 39 (2): 53–78.

Lakdawalla, Darius N., and Neeraj Sood. 2012. "Incentives to Innovate." In *The Oxford Handbook of the Economics of the Biopharmaceutical Industry*, edited by Patricia M. Danzon and Sean Nicholson. Oxford: Oxford University Press.

Lakdawalla, Darius N., Jalpa A. Doshi, Louis P. Garrison Jr., Charles E. Phelps, Anirban Basu, and Patricia M. Danzon. 2018. "Defining Elements of Value in Health Care—A Health Economics Approach: An ISPOR Special Task Force Report [3]." *Value in Health* 21 (2): 131–39.

Moreno, Santiago G., and Jordan A. Ray. 2016. "The Value of Innovation under Value-Based Pricing." *Journal of Market Access & Health Policy* 4 (1).

Nordhaus, William D. 1969. "An Economic Theory of Technological Change." *American Economic Review* 59 (2): 18–28.

Peltzman, Sam. 1976. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19: 211–40.

Polk, Andreas, Armin Schmutzler, and Andreas Müller. 2014. "Lobbying and the Power of Multinational Firms." *European Journal of Political Economy* 36: 209–27.

Rand, Lisa Z., and Aaron S. Kesselheim. 2021. "International Reference Pricing for Prescription Drugs in the United States: Administrative Limitations and Collateral Effects." *Value in Health* 24 (4): 473–76.

Scannell, Jack W., Alex Blanckley, Helen Boldon, and Brian Warrington. 2012. "Diagnosing the Decline in Pharmaceutical R&D Efficiency." *Nature Reviews Drug Discovery* 11 (3): 191–200.

Woods, Ben, Jon Lomas, Mark Sculpher, Helen Weatherly, and Karl Claxton. 2024. "Achieving Dynamic Efficiency in Pharmaceutical Innovation: Identifying the Optimal Share of Value and Payments Required." *Health Economics* 33 (4): 804–19.

Wouters, Olivier J., Martin McKee, and Jeroen Luyten. 2020. "Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009–2018." *JAMA* 323 (9): 844–53. — — —. 2022. "Erratum." *JAMA* 328 (11): 1110–11.

Figure 1: Company-specific trends in (a) scaled cumulative pharmaceutical R&D expenditures, $I_{m,y}$ in Eq (8), and (b) probability of regulatory success, $\pi_{m,y}$ in Eq (12).

Notes: Scaled cumulative pharmaceutical R&D expenditures subtract the observed first year of R&D expenditures from the cumulative expenditures going forward, so every company starts with zero costs. The probability of regulatory success is determined by the cumulative number of approvals each company receives, relative to the cumulative number of assets invested over time. It reflects the overall performance of a company as it continues to invest in research and development.

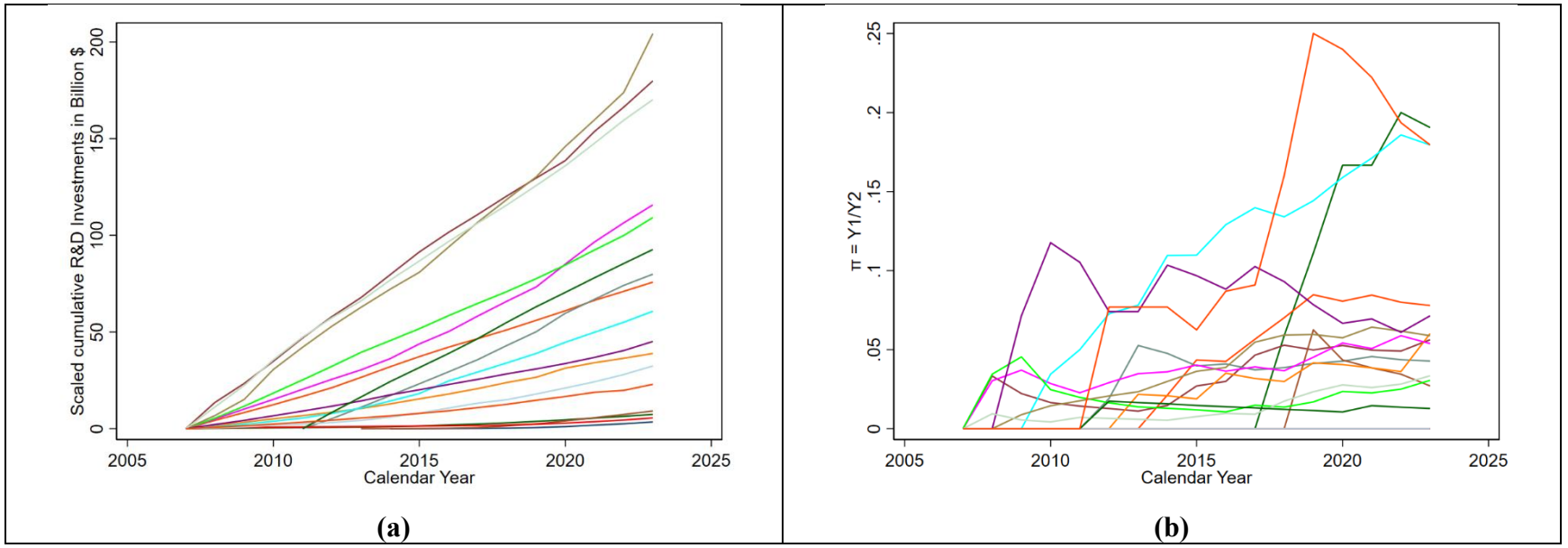


Figure 2: Estimated production function for the biomedical industry in the US

Notes: This estimates an aggregate industry-level production function for the probability of regulatory success as a function of R&D investments

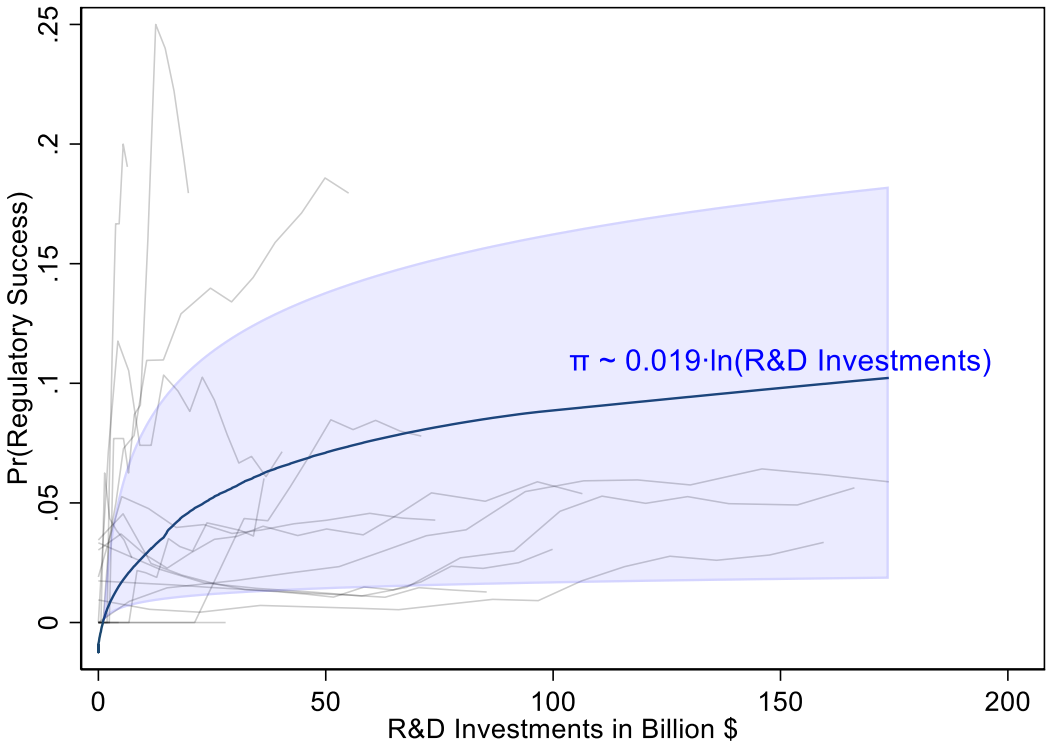


Figure 3: Optimal portion of surplus transfer to innovators conditioned on the relative weight (Lambda, λ in Eq(6) that the social planner puts on consumer surplus in their objective function.

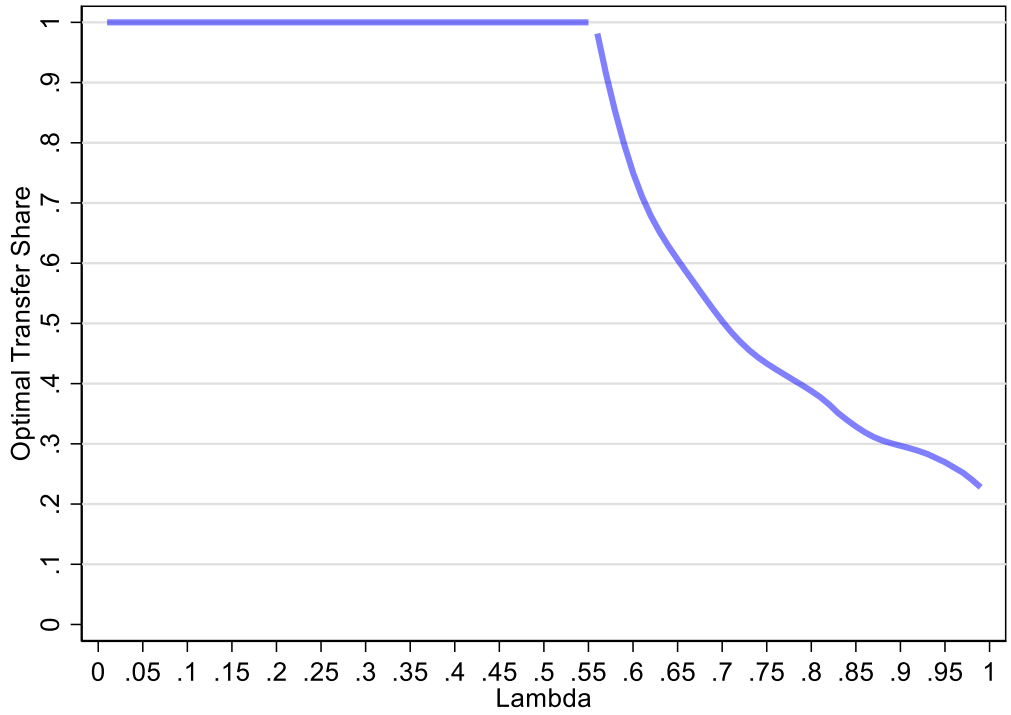
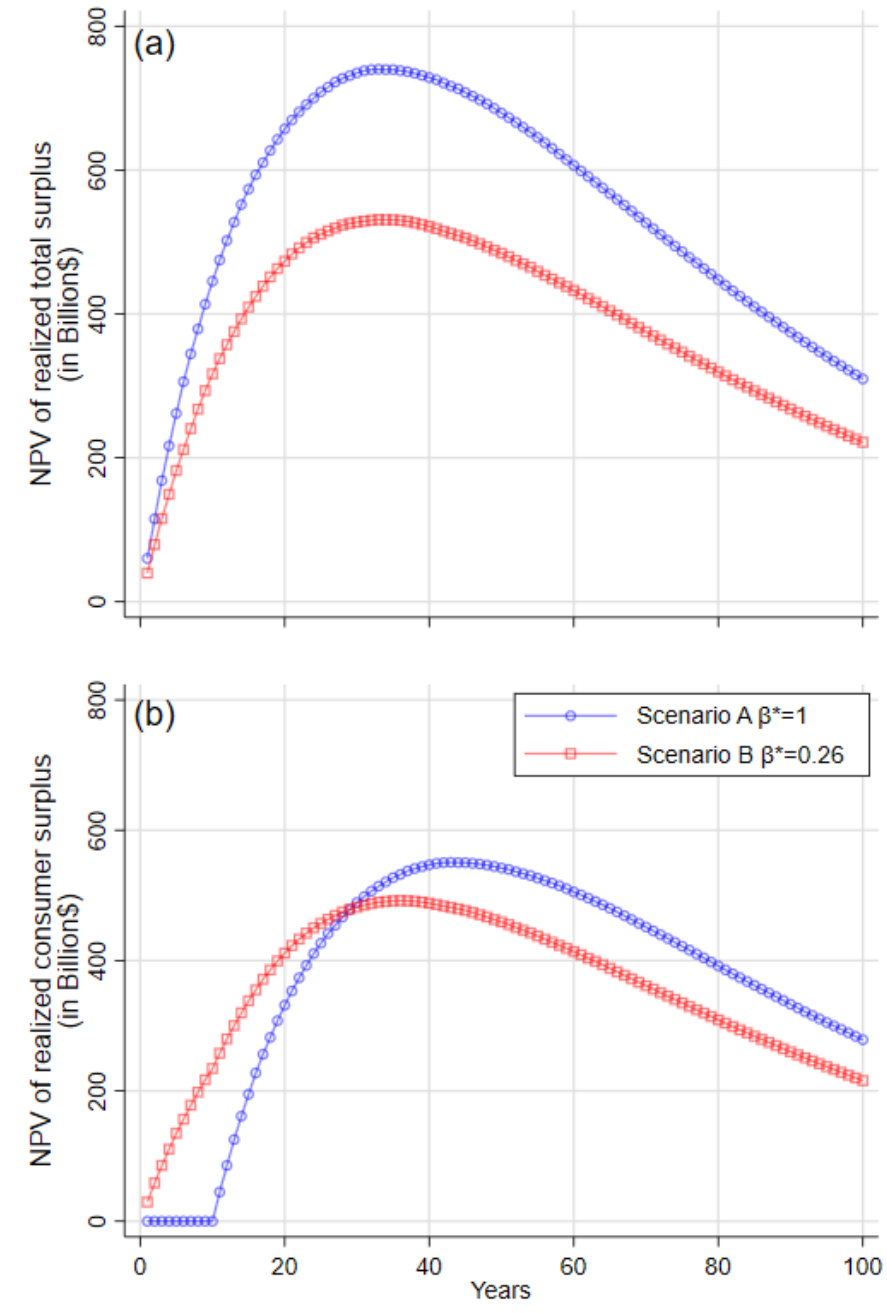
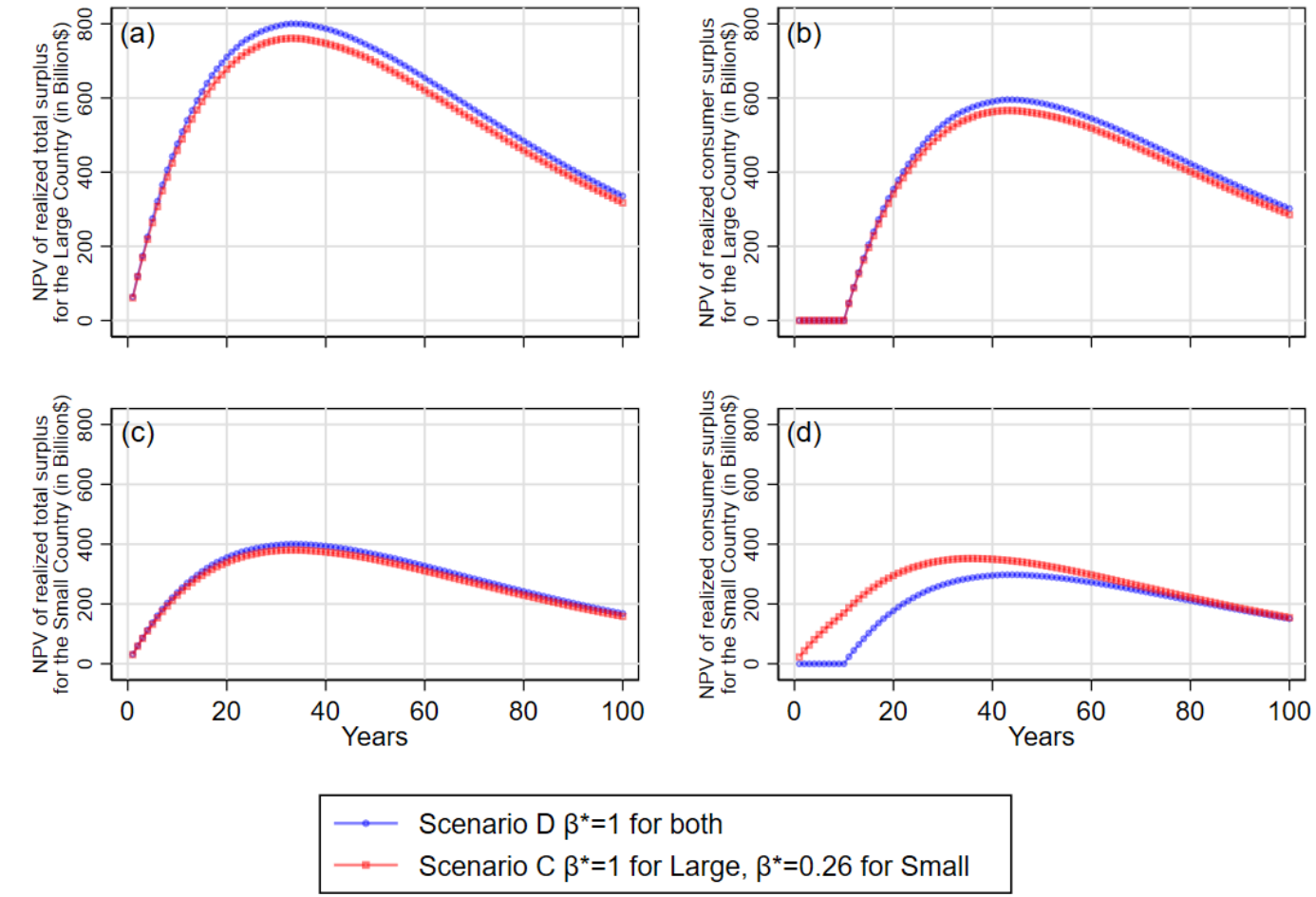


Figure 4: Net Present Value (NPV) of the realized (a) total and (b) consumer surplus by chronological year under innovation policy A (Only one large country exists, and it maximizes total surplus, i.e. $\beta^* = 1$) and policy B (Only one large country exists, and it maximizes consumer surplus, i.e. $\beta^* = 0.26$)



Discounted at 3%

Figure 5: Net present value (NPV) of realized (a) total and (b) consumer surplus by chronological year for the large country and (c) total and (d) consumer surplus for the small country under innovation policy C (A large and a small country coexist. The large country maximizes total surplus, i.e. $\beta^* = 1$. But the small country maximizes consumer surplus, i.e. $\beta^* = 0.26$) and policy D (A large and a small country coexist. Both countries maximize total surplus, i.e. $\beta^* = 1$)



Discounted at 3%