The Effect of Market Size on the Cost and Quality of Pharmaceutical Innovation

Alice Chen
University of Southern California

Partha Deb
City University of New York and NBER

Darius Lakdawalla
University of Southern California and NBER

October 11, 2024

Abstract. We develop and test the implications of upward-sloping innovation supply for the cost and quality of new innovations that arrive when pharmaceutical demand expands. When demand increases in the presence of upward-sloping supply, innovators become more willing to bet on drug candidates with higher risks of failure. Empirically, we find that exogenous demand growth triggered by Medicare Part D increased the risk of drug candidate approval failures by 7.7% a year. As a result, the marginal drug induced by Part D demand growth cost $530 million (or 24%) more to launch than the average drug. The increased cost bought additional quality, as marginal innovations were 12% more likely to represent major clinical improvement. Each $1 billion of additional pharmaceutical revenue leads to 1.6 more drugs bringing significant, major, or moderate clinical improvement and 0.9 fewer drugs with no clinical improvement.

I. Introduction

In 1959, Senator Estes Kefauver convened a landmark series of hearings to explore pricing and competition in the pharmaceutical industry. The Senate Subcommittee on Anti-Trust and Monopoly performed its work for nine months and ultimately released dueling reports. The Majority Report argued that US pharmaceutical prices were excessive by international standards, abetted by anti-competitive behavior, and unjustified by the incremental clinical innovation produced by the industry. In contrast, the Subcommittee's Minority Report argued that prices were reasonable reflections of the risk undertaken by the industry and the high level of scientific innovation it produced for American consumers (Morgan and Allison 1964).

The passage of 60 years has done little to resolve the debates that played out in the Kefauver Hearings. These same themes have been struck by policymakers debating the merits of the 1984 Hatch-Waxman Act, the 2003 Medicare Modernization Act, and even the 2022 Inflation Reduction Act. Sachs, Adler, and Frank (2023) point out at least one reason why these controversies persist: Despite ample evidence that higher prices and profits lead to greater quantities of pharmaceutical innovation, its resulting clinical value to patients remains a matter of debate. Indeed, proponents of stricter industry regulation have argued that pharmaceutical firms pursue so-called "me-too" pharmaceutical innovation that provides limited benefit to patients (Angell 2004), on the premise that scientifically modest innovations are cheaper to develop and launch (Aronson and Green 2020). If true, these concerns imply that enhancing financial rewards for innovators encourage new drugs that are cheap to develop, profitable to sell, but clinically incremental for patients. On the other side of this debate, proponents of stimulating pharmaceutical innovation point to robust economic evidence that financial rewards stimulate the quantity of innovation (Acemoglu and Linn 2004; Blume-Kohout

and Sood 2013; Dubois et al. 2015; Finkelstein 2004), and that the *average* innovation has produced substantial benefits to patients (Philipson and Jena 2006), to conclude that stimulating innovation improves health and life expectancy (Philipson and Durie 2021). While these economic studies remain salient, they do not shed light on the value of the marginal new drug, which remains hotly contested in debates over the wisdom of drug price controls.

To help unravel this controversy, we propose and test a hypothesis linking growth in market size to the quality and costs of the marginal innovations that it induces. Our theoretical conjecture follows Myers and Pauly (2019). They apply the Ricardian theory of productivity to the pharmaceutical industry and show that demand growth causes firms to reach for costlier innovations on the margin, in much the same way that David Ricardo imagined growth in the demand for food pressing into service less fertile and productive farmland. Myers' and Pauly's results suggest an upward-sloping supply of pharmaceutical innovation, where the marginal drug is more expensive to launch than the average drug that came before it.

[Insert Figure 1 Here]

Other stylized facts fit the finding of demand rising along an upward-sloping supply curve for innovation. Rising costs per new drug launch have coincided with growth in the prices and real revenues of newly launched drugs. Using Medical Expenditure Panel Survey (MEPS) data, Figure 1 reports that real prices are 71.6% higher for the average new drug launched in 2015 compared to its counterpart launched in 1996.[1]

[Insert Figure 2 Here]

[1] As we discuss in Section IV, we supplement the MEPS data with approval dates from the National Drug Code Directory. MEPS data do not account for rebates paid by manufacturers to downstream payers. Net prices, inclusive of rebates, show similar upward trends, although data are only available from 2007 onwards, in the SSR Health database.

The average inflation-adjusted revenues accruing to each new drug have also been weakly rising over time. Figure 2 reports average three-year revenues per new drug, measured four to six years after the date of approval, also based on MEPS data. In addition, the health economics literature provides reasons to believe in rising pharmaceutical demand over time. For example, Phelps and Cinatl (2021) point out that the demand for quality-adjusted life-years (QALYs) grows in proportion to per capita consumption (net of healthcare spending).[2]

Like Ricardo's conjecture for farmland, we conjecture that demand growth encourages pharmaceutical manufacturers to place bets on riskier (more speculative) drug candidates on the margin, because stronger demand offsets the costs of lower approval probabilities. If demand growth leads to greater R&D failure rates, it will result in consistently rising R&D costs per drug launch, drug revenues, drug candidates, and drug approvals. We test our hypothesis using Medicare Part D as an exogenous shock to demand, the same experiment used by a number of prior papers studying incentives for pharmaceutical innovation (Bennette et al. 2019; Dranove, Garthwaite, and Hermosilla 2020; Krieger, Li, and Papanikolaou 2021; Blume-Kohout and Sood 2013).

Contrary to the emphasis on "me-too" drugs, we show that the marginal innovations stimulated by Part D suffered higher failure rates, and that these marginal drugs launched at higher costs as a result. We estimate that the marginal drug candidate induced by a 50% increase in market size fails at a rate of 98.4%, substantially higher than the average pre-market expansion failure rate of 91.5%. Our estimates imply that

[2] Specifically, they define H as an index of health-related quality of life improvement and C as consumption net of health spending, where utility is given by $U(C, H)$. In this framework, the willingness to pay for a unit change in quality-adjusted life-years is given by $\frac{C}{\omega_C}$, where $\omega_C \equiv \frac{U'(c)c}{U(C)}$, the elasticity of utility with respect to consumption. The latter is a constant preference parameter under constant relative risk-aversion (CRRA) preferences.

for a 50% increase in the expected net present value of revenues, the marginal drug costs between $0.40 billion to $2.54 billion more to launch than the average drug, depending on which phase-specific R&D cost estimates from the literature are applied. In addition, by imposing a bit of additional structure on our empirical framework, we recover a lower bound on the Ricardian rents flowing to the pharmaceutical industry. We estimate these are at least $0.88 billion for each inframarginal drug launch. This amounts to a marginal R&D cost share of around 34%, which is broadly consistent with the gross margin of pharmaceuticals net of production, distribution, and marketing costs estimated in the literature.

Our evidence also shows that higher R&D failure rates and costs translate into higher clinical quality on the margin. We also show that the marginal drug launch induced by Part D was more likely to produce major clinical benefits as the drugs launched before the insurance expansion. Following prior research (Krieger, Li, and Papanikolaou 2021; Kyle and Williams 2017), we rely on France's Haute Autorité de Santé (HAS) agency's assessment of clinical benefit for newly launched drugs. The share of drugs that HAS rated as producing moderate to major improvement increases by 11.6% after Part D. An extra $1 billion of market size is estimated to produce 1.6 more drugs with major clinical benefits and 0.9 fewer drugs with no clinical benefit, as defined using the HAS criteria of benefit.

Our paper relates to a significant prior literature on the economics of pharmaceutical innovation. Nordhaus (1969) proposed a canonical economic model of innovation, implying that greater financial rewards stimulate more inventive effort and output. Subsequent empirical research validated this prediction within the context of the pharmaceutical industry (Acemoglu and Linn 2004; Blume-Kohout and Sood 2013; Dubois et al. 2015; Finkelstein 2004). On the subject of innovation quality, Bennette et al. (2019)

study the particular context of oral chemotherapy, showing that greater rewards lead to innovation that increases survival, but by less than its predecessors. This paper contributes to the literature by studying whether increases in expected rewards for tomorrow's innovations influence innovation quality, risk, and development costs. This addresses the longstanding policy question of whether higher expected rewards for innovators yield clinically impactful drugs for patients. Krieger, Li, and Papanikolaou (2021) have previously shown that positive shocks to internal capital, for instance due to increases in profits from *currently* marketed products, cause firms to invest in and launch riskier, more clinically beneficial innovations. However, it is not clear if expected future rewards have the same effect, outside the context of oral chemotherapy. This distinction matters for policy, because expected rewards remain a more finely calibrated policy lever than current cash flows for achieving particular societal goals. For example, public policies can target expected rewards for particular kinds of drugs (e.g., rare disease drugs) to stimulate innovation of that kind, but policies stimulating cash flow alone may produce innovation in any number of different disease areas or contexts.

Our paper is organized as follows. Section II proposes and develops a stylized model of upward-sloping innovation supply. Section III outlines our empirical approach, which uses Part D as an exogenous shock to market demand. Section IV describes our sources of data. Section V presents our results, and Section E concludes.

II. Theoretical Framework for Estimation

We develop a simple, stylized model of growth in demand for pharmaceuticals in the presence of upward-sloping innovation supply. The model illustrates the key assumptions that underlie our predictions, and it also provides a blueprint for our empirical analysis.

A. The Innovation Environment

Basic scientific discoveries create a flow of potential new drug candidates that arrive each period. For a given drug class c and time t, these potential drug candidates lie on a unit interval and can be indexed by $i_{ct} \in [0,1]$. Pharmaceutical companies must decide which of these molecules to pursue in development. Notice that we analyze innovators' decisions about a per period flow of new drug candidates, not a permanently fixed stock. Each period a new set of drug candidates arrives via basic scientific discoveries, and a new set of decisions is made.

Suppose that the innovator expects to earn the net present value of profits, $\pi(i_{ct}; m_{ct})$, conditional on candidate i in drug class c successfully launching at time t. Profits depend in part on m_{ct}, the size of the market for drugs in class c that is expected to prevail at time t. All other determinants of expected profits (or revenues) are held constant in our theoretical framework, including the distribution of clinical benefit and the associated "quality" of potential drug candidates arriving each period. We can think of $D(i; m) \equiv \pi(i; m)$ as the "demand curve" for investigational new drug candidates, because it traces out the willingness to pay curve. It is not *a priori* obvious whether the willingness to pay rises or falls with i. Thus, the "demand curve" in this case can slope upward, making the supply-demand analogy imperfect. However, the intuition associated with the supply-demand framework still proves illustrative, even though we preserve throughout the possibility that the demand curve slopes upward or downward.

On the "supply-side" of the innovation decision, define the cost of researching and developing drug candidate i as $\delta(i)$, the risk of failure as $\phi(i)$, and the corresponding chance of success as $\sigma(i) \equiv 1 - \phi(i)$. The expected cost of bringing candidate i to market is $\frac{\delta(i)}{1-\phi(i)} = \frac{\delta(i)}{\sigma(i)}$. To illustrate, suppose the risk of failure is 90%, the chance of

success 10%, and the cost of developing each candidate \$100m. In this example, it takes an average of 10 candidates to produce one successful launch and an average of \$1 billion to bring one drug to market, or $\frac{\$100\text{m}}{0.1} = \$1\text{bn}$. Since firms need to be compensated for the expected cost of bringing their drugs to market, we think of $\frac{\delta(i)}{\sigma(\text{i})}$ as the "innovation supply curve," defined by $S(i) \equiv \frac{\delta(i)}{\sigma(\text{i})}$ and yielding the "willingness to accept" curve for the firm.

In our empirical analysis, we will treat δ as fixed across drugs but allow the failure and success probabilities, $\phi(\text{i})$ and $\sigma(i)$, to vary, for both practical and theoretical reasons. From a practical standpoint, there is no publicly available data on drug- or class-specific development cost.[3] Therefore, it is not possible to directly estimate the effect of exogenous factors like market size growth on development costs. Moreover, from a theoretical standpoint, variation in the risk of failure plays an outsized role in the cost of launch. For instance, observe that increasing the risk of failure from 90% to 95% doubles the number of drug candidates needed for one launch, from 10 to 20, and thus doubles the required launch cost. Since empirical failure risk is very high in the pharmaceutical industry, this provides a great deal of empirical leverage to the risk of failure in determining launch costs.

B. Identifying Innovation Supply and Demand Parameters

Next, we define several ways of empirically measuring both the riskiness and costliness of drug candidates, and we explain how they relate to the notional demand for and supply of innovation.

---

[3] There is some data on changes over time in aggregate development costs, and we will use these later (DiMasi, Grabowski, and Hansen 2016; DiMasi, Hansen, and Grabowski 2003). There are also class-specific R&D data published by the industry trade group PhRMA, but reporting ceased in 2000, prior to the passage of Medicare Part D (Myers and Pauly 2019).

The representative pharmaceutical firm will decide to develop new drug candidates until the expected cost of developing the next candidate exceeds the expected benefit the firm can earn from doing so. The firm's expected benefit from launching drug i will be $\sigma(i)\pi(i;m)$. For the marginal drug, i^*, the expected profit from pursuing the drug candidate, $\sigma(i^*)\pi(i^*;m)$, equals the fixed cost of development (δ), or

$$E\pi_{i^*} \equiv [\sigma(i^*)]\pi(i^*;m) = \delta \quad (1)$$

Notice that we can rewrite this condition in terms of equality between the notional supply of and demand for innovation at the marginal drug candidate, i^*:

$$D(i^*;m) \equiv \pi(i^*;m) = \frac{\delta}{[\sigma(i^*)]} \equiv S(i^*) \quad (2)$$

The term $\frac{\delta}{\sigma(i)}$ is the marginal cost of launching the i^{th} drug candidate. Therefore, we have the standard result that the supply curve for drug launches coincides with the marginal cost of drug launches. With this heuristic supply-demand framework in place, we can define a few conceptual terms for which we will provide empirical analogues later.

1. Defining Innovation Parameters

a) Marginal Launch Cost

Define $MLC(i)$ as the "marginal launch cost" for the i^{th} drug candidate, where $MLC(i) \equiv \frac{\delta}{\sigma(i)} \equiv S(i), \forall\, i.$

b) Average Launch Cost

Analogously, define the "average cost of launching i drug candidates" as $ALC(i) \equiv \frac{\delta}{\overline{\sigma(i)}}$, where the average success probability is defined as $\overline{\sigma(i)} \equiv \frac{1}{i}\int_0^i \sigma(j)dj$ and analogously for the average risk of failure, $\overline{\phi(i)} \equiv \frac{1}{i}\int_0^i \phi(j)dj$.

c) Number Needed to Develop

The reciprocal of the average launch cost yields the "number needed to develop," or the expected number of drug candidates that must be researched to generate one new drug, as in $\overline{NND}(i) \equiv \frac{1}{\overline{\sigma(i)}}$.

Since real-world success probabilities are relatively low for pharmaceuticals, even small changes in these success probabilities can substantially change the number needed to develop. For instance, if success probability falls from 15% to 10%, just five percentage points, the number needed to develop rises by 50%.

d) Candidates, Failures, and Launches

The equilibrium number of drug "candidates" is given by i^*. Define the equilibrium number of "failures" as $\Phi(i^*) \equiv \int_0^{i^*} \phi(j)dj$, and conversely define the number of "launches" as $\ell(i^*) \equiv (i^* - \Phi(i^*))$.

e) Marginal failure risk

The "marginal risk of failure" is defined as the failure rate for the marginal drug candidate, or $\phi(i^*)$.

f) Average failure risk

The "average risk of failure" is given by $(\overline{\phi}(i^*) \equiv \frac{\Phi(i^*)}{i^*})$.

2. Identifying Innovation Parameters

Our empirical strategy assumes that exogenous shifts in market size, m, affect only post-launch profits, π. As a result, shifts in m allow us to recover the effect of changes in profits, π, on failures and successes. To pursue this point further, notice that $\left(\frac{\partial}{\partial \pi}\Phi(i^*)\right) = \frac{\frac{\partial}{\partial m}\Phi(i^*)}{\frac{d\pi}{dm}}$. Therefore, so long as we can identify the effect of market size on failures, $\left(\frac{\partial}{\partial m}\Phi(i^*)\right)$, and its effect on profits, $\frac{d\pi}{dm}$, we can then identify the effect of changes in profits on failures, $\left(\frac{\partial}{\partial \pi}\Phi(i^*)\right)$. By similar logic, if we can identify the effect

of market size on the number of drug candidates, $\frac{\partial}{\partial m} i^*$, we can identify the effect of changes in profits on i^*, because $\left(\frac{\partial i^*}{\partial \pi}\right) = \frac{\frac{\partial}{\partial m} i^*}{\frac{d\pi}{dm}}$.

The approaches above allow us to recover $\left(\frac{\partial}{\partial \pi} \Phi(i^*)\right)$ and $\left(\frac{\partial i^*}{\partial \pi}\right)$. With these in hand, we can identify the failure rate for the marginal drug candidate, $\phi(i^*)$. Since the definition of $\Phi(i^*)$ implies that $\frac{\partial}{\partial m} \Phi(i^*) = \frac{\partial i^*}{\partial m} \phi(i^*)$, it follows that $\phi(i^*) = \frac{\frac{\partial}{\partial \pi} \Phi(i^*)}{\frac{\partial i^*}{\partial \pi}}$, a ratio of two estimable parameters. Intuitively, the marginal failure rate equals the marginal failures induced by a demand shift divided by the marginal number of drug candidates induced.

To estimate launch costs in our framework, we make use of prior literature that estimates development costs, δ, in aggregate (DiMasi, Hansen, and Grabowski 2003; Adams and Brantner 2006; DiMasi, Grabowski, and Hansen 2016; Sertkaya et al. 2024). Given an estimate of δ, we can estimate the equilibrium marginal launch cost, $MLC \equiv \frac{\delta}{1-\phi(i^*)}$, and the average launch cost, $ALC \equiv \frac{\delta}{1-\bar{\phi}(i^*)}$. Moreover, note that MLC also equals $\pi(i^*; m)$, the net present value of profits that accrues from the marginal drug candidate.

C. Testable Implications of Upward-Sloping Innovation Supply

Recovering the parameters defined above allows us to test several implications of the upward-sloping supply hypothesis. The hypothesis itself matches two empirical facts: 1) Growth in market size leads to more drug approvals; and 2) Growth in market size also increases the risk of failure. The first fact has been demonstrated by prior literature (Acemoglu and Linn 2004; Blume-Kohout and Sood 2013; Dubois et al. 2015). The second we will demonstrate in our empirical analysis.

To comport with these facts, two things must be true about the notional supply and demand curves we have described:

$$MLC'(i) \equiv S'(i) > 0 \qquad \text{(Condition 1)}$$

$$S'(i) > D_i(i;m) \equiv \pi_i(i;m) \qquad \text{(Condition 2)}$$

Condition (1) requires that supply must be strictly upward-sloping. Holding development costs constant, therefore, success probabilities must be lower for the marginal drug than for the average drug. Condition (2) accommodates both upward- and downward-sloping demand cases. It requires only that the slope of the demand function be less than the slope of the supply function. Prior empirical research supports Condition (2), because were this condition to fail, growth in market size would lead to fewer drug candidates and approvals. This logic unfolds in the following implication.

Implication 1: *If Condition (2) holds, then growth in market size leads to more drug candidates. Specifically, if* $\boldsymbol{D_i(i;m) < S_i(i)}$ *then* $\boldsymbol{\frac{di^*}{dm} > 0}$.

We prove this result in Appendix A: Mathematical Derivations, but Figure 3 illustrates the intuition by depicting notional supply and demand curves that match these two conditions.

[Insert Figure 3 Here]

In the figure, the supply curve, i.e., the marginal launch cost curve, $MLC(i)$, slopes upward to match Condition (1). This also implies that $ALC(i)$ is upward-sloping. The downward slope of the notional demand curve, $\pi(i;m)$, satisfies Condition (2), but it bears repeating that upward-sloping demand curves remain possible in our mathematical derivations here. The figure illustrates the effect of exogenous growth in market size, m, that shifts the post-launch profits curve from π_0 up to π_1. Holding development costs, δ, constant, the number of drug candidates rises from i_0^* to i_1^*. At the same time, demand-driven expansions in revenue lead to higher marginal launch costs, so long as marginal launch costs slope upward with the number of drug candidates. Intuitively,

with upward-sloping supply, developers press lower probability drug candidates into development when demand rises.

Condition (1), or upward-sloping supply, ensures that productivity and success rates will fall with demand expansion.

Implication 2: If *Condition (1) holds, the average risk of failure rises with market size, or* $\frac{d\overline{\phi(i^*)}}{di^*}\frac{di^*}{dm} > 0$.

To prove the result, note that Implication 1 implies $\frac{\mathrm{di}^*}{\mathrm{dm}} > 0$. By Condition (1), ($\frac{\mathrm{d}\overline{\phi(i^*)}}{\mathrm{di}^*} > 0$). Therefore, the result follows. This in turn immediately yields a corollary characterizing average launch costs and marginal launch costs.

Corollary 2.1: *Average launch cost rises with market size, or* $\frac{dALC(i^*(\pi))}{dm} \equiv \frac{\delta}{[1-\phi(i^*)]^2}\frac{d\overline{\phi(i^*)}}{di^*}\frac{di^*}{dm} > 0$.

When drug companies turn to riskier drug candidates, the average cost of drug development also rises. Since market size increases average failure risk, it also increases the number needed to develop, as in the following corollary.

Corollary 2.2: "*Number needed to develop" (NND) rises with post-launch profits, or* $\frac{d\overline{NND}(i^*)}{di^*}\frac{di^*}{dm} > 0$.

Recall from Section B.2 above that our framework permits estimation of failure risk for the marginal drug, using exogenous shifts in market size. This empirical parameter enables us to test the following implication concerning rising marginal launch costs.

Implication 3: *The marginal failure risk exceeds the average failure risk, or* $\phi(i^*) > \overline{\phi(i^*)}$*, because average failure risk rises with market size.*

The proof follows directly from Condition 1. An immediate corollary follows.

Corollary 3.1: *Provided development costs are constant or nondecreasing, the marginal launch cost will exceed the average launch cost.*

The proof follows because marginal and average launch costs are simply the reciprocal of marginal and average success probabilities, multiplied by development costs, δ.

III. Empirical Approach

Our goal is to identify how shifts in the demand for pharmaceuticals affect new drug approvals, new drug candidates, candidate failures, R&D failure rates, launch costs for the marginal and average drug candidates, and quality of the marginal drug launched. Our identification strategy uses a difference-in-difference design that exploits the implementation of the Medicare Part D legislation as an exogenous shock to demand that is plausibly unrelated to the supply of pharmaceutical innovation. Introduced via the Medicare Modernization Act on December 8, 2003, and formally implemented January 1, 2006, Medicare Part D established outpatient drug benefits for Medicare beneficiaries.[4] Prior to 2006, traditional (fee-for-service) Medicare did not provide outpatient drug coverage to beneficiaries.

[Insert Figure 4 Here]

Figure 4 provides suggestive evidence that Part D affected R&D failure rates and hence launch costs. The figure plots the raw, unadjusted number of Phase I drug candidates needed to approve one drug within eight years. Prior to Part D, approximately one of seven Phase I drug candidates ended up being approved. By 2011, however, this success

[4] The law passed the House vote on June 27th, 2003, and the Senate vote on November 25th, 2003.

rate had fallen to roughly one out of eleven. To obtain the second source of variation for the difference-in-difference design, we follow prior literature in assuming that Part D produced differential effects on demand in drug classes with high and low shares of elderly users (Blume-Kohout and Sood 2013).

For simplicity and following the prior literature (Acemoglu and Linn 2004; Dubois et al. 2015), we abstract from marginal costs of production and measure post-launch profits using the net present value of revenues that would be earned if the drug is launched. To mitigate bias due to the use of revenues, we focus our attention on the effects of percentage, rather than absolute dollar, changes in revenues. Thus, as long as post-launch percentage profit margins are uncorrelated with our error terms, our measurement approach will produce reliable results.[5] We compute effects on failure risks and development costs in three steps: 1) predict life-cycle revenue profiles for each drug; 2) predict expected net present value of revenues in each drug class; and 3) predict drug candidates, approvals, and failures as a function of expected net present value and other controls. We then estimate the effect of Part D on the quality of innovation.

A. Predicting Drug-Level Revenue Profiles

In the first step, we specify a model of revenues at the drug-year level, as a function of a dummy variable demarcating Part D implementation (year $\geq$ 2006), a continuous variable measuring the Medicare market share (MMS) in the drug's class, and an interaction between the Part D indicator and MMS. To eliminate the possibility of reverse causality, we calculate Medicare market share from 2001 to 2003, before Part D was passed and implemented.[6] We also exclude drugs launched after the passage of Part D, since

[5] Note that profits equal revenues multiplied by the profit margin. Thus, percentage changes in revenues equal the percentage growth in profits minus the percentage growth in profit margins.

[6] Our results are robust to the calculation of Medicare market share using different pre-period years.

these launch decisions are endogenous to Part D itself. The estimation additionally includes calendar year (ϕ_t), time-since-approval ($\phi_{d(t)}$), and drug class (ϕ_c) fixed-effects. In addition, we control for additional covariates (X_{dt}), including an indicator for Medicare Part D's passage (year $\geq$ 2003), an interaction of the passage indicator with the Medicare market share (MMS) in the drug class, and an indicator for generic entry.[7] We assume that our covariates influence the percentage growth in drug revenue in a linear fashion. For drug d in class c at calendar year t, we assume that the relationship between revenues and covariates can be represented by an exponential conditional mean function, i.e.,

$$E(R_{dct}) = \exp\begin{pmatrix}\alpha_0 + \alpha_1 \text{Part D}_t + \alpha_2 \text{Part D}_t * \text{MMS}_c + \\ \alpha_X X_{dt} + \phi_c + \phi_t + \phi_{d(t)}\end{pmatrix} \quad (3)$$

The parameters of this conditional mean function are estimated using a Generalized Linear Model (GLM) with a logarithmic link function. GLM is a generalization of ordinary least squares (OLS) regression that allows the prediction error to be drawn from any distribution in the exponential family.[8] GLM is also convenient in our setting, because it directly produces a consistent estimate of $E(R_{dct})$. In contrast, OLS estimation of a model with log revenues as the dependent variable would produce an estimate of $E(\ln R_{dct})$ that would require a non-trivial transformation to obtain estimates of $E(R_{dct})$. We assume that the revenue function follows a Gamma distribution, because we observe that its distribution is considerably right skewed. Note that the gamma distribution is also the leading distributional choice for health expenditures (Deb, Norton, and Manning 2017). GLM parameter estimates feature a powerful robustness property:

---

[7] In an alternative model, we included an additional indicator for years greater than 2010—which to the filling of the Part D donut hole—and its interaction with MMS. While the donut hole created an additive increase in revenues for drugs with higher MMS, its effect was not statistically significant.

[8] Since the analyst must specify one of these distributions, we tested the full range of distributional assumptions.

they are consistent as long as the conditional mean is correctly specified, i.e., even if the distribution of the outcome is incorrectly specified.

B. Predicting Class-Level Expected Revenues

With estimated revenue profiles for our drugs in hand, we turn to the second step of our estimation problem. The theory illustrates that innovators act on their expectations about market size when making development decisions. This expectation is unobservable, and we need to employ some assumptions about how innovators form their beliefs. Following prior research and our theoretical framework, we define drug classes as distinct markets (Dubois et al. 2015). In addition, we assume that expected growth in the expected net present value (ENPV) of previously launched drugs in a class serves as a proxy for exogenous expected growth in market size.

To construct the ENPV, we exploit the drug life-cycle revenue profiles to estimate the expected revenue for each drug during the first 11 years of the drug's life cycle since launch. The first step is to estimate fitted values from Equation (3), holding calendar year constant at 2001 and time-since-approval constant at 5. (Our results are robust to different calendar year and time-since-approval baselines.) To predict revenues that would have accrued in the absence of Part D, denoted by $\widehat{R^0_{dt}}$, we additionally turn off the Part D treatment effect by setting $\widehat{\alpha_2}$ in Equation (3) to zero. In contrast, to predict revenues under Part D, denoted by $\widehat{R^1_{dt}}$, we rely on $\widehat{\alpha_2}$ to identify the plausibly exogenous changes in revenues due to the differential impact of Part D by Medicare market share (i.e., the treatment effect).

Next, we calculate the drug-level ENPV by summing discounted predicted revenues with and without Part D. We define these two streams of ENPV by $\widehat{ENPV^1_d} \equiv \sum_{j=0}^{10} \frac{1}{(1+r)^j} \widehat{R^I_{dj}}$, where $\widehat{R^I_{dj}}$ equals predicted revenues during the first 10 years of drug's life cycle had Part D not occurred ($i = 0$) or with Part D ($i = 1$). We set the discount rate r

at 10%, which was the cost of capital for the biotechnology sector in January 2005.[9] Finally, we calculate the class-level ENPV by summing the drug-level ENPV over all drugs present in class c in year t.[10] Specifically, we define class-level predicted ENPV by:

$$\widehat{ENPV}_{ct} \equiv \begin{cases} \sum_{d \in \{c_t\}} \widehat{ENPV}_d^0 & \text{if } t \leq 2003 \\ \sum_{d \in \{c_t\}} \widehat{ENPV}_d^1 & \text{if } t \geq 2004 \end{cases}$$

C. Predicting Drug Candidates, Failures, and Approvals

In the next step of our empirical analysis, we are interested in how plausibly exogenous changes in market size, i.e., ENPV in the relevant class, affect trial failures, drug approvals, and drug candidates in that class. Define Y_{ct} as a development outcome in class c for drug candidates that went into Phase I development in year t. We consider the following development outcomes, Y, all of which are measured as count data: Phase I trials, Phase II trials, Phase III trials, drug approvals, failures from Phases I to II, failures from Phase II to III, failures from Phase III to approval, and failures from Phase I to approval.[11] Note that our definition of t employs the "cohort" approach illustrated in our stylized theoretical model, grouping together all drugs that entered development in year t. Here, we assume that market size at time t influences the cohort of drugs entering development in this year and thus influences all future development outcomes for that cohort. From the date of cohort entry, progression to future phases may be censored by data availability. We create uniformed censoring by defining our outcome variables as trials occurring within 12 years of the Phase I start date: if the next phases had

[9] Archived data from Aswath Damodaran: https://web.archive.org/web/20050429035437/https://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/wacc.html.

[10] Recall that the subset of drugs we consider are only those that were approved prior to Part D.

[11] Phase I trials are early trials testing for drug safety. Phase II trials are the first efficacy trials in small numbers of patients. Phase III trials are large scale trials for registration data.

not occurred within 12 years of the Phase I start date, we counted the drug has having failed to progress. We test the robustness of our results to different durations of drug development and discuss the distribution of trial durations in Section IV.

Since our outcome variables are counts, we model our development outcomes Y_{ct} using a Poisson regression (Dubois et al. 2015; Duggan and Scott Morton 2010). The Poisson regression also has desirable robustness properties; parameter estimates are consistent as long as the conditional mean is correctly specified. Defining X_c as a set of class-level controls, including MMS, share of generic entry, and average years since drug approval, we estimate the relationship between drug development and Part-D induced changes in market size with:

$$E(\Delta Y_{ct}) = \exp\left(\beta_0 + \beta_1 \Delta \ln\left(\widehat{ENPV}_{ct}\right) + \beta_2 X_c + ln(\overline{Y}_c)\right) \quad (4)$$

By controlling for the average of the outcome variable, $\overline{Y}_c$, as an exposure, we allow counts of drug development to be larger when there is more activity in the drug class (i.e., when the class is exposed to more trials and failures on average). This is a Mundlak-type correction for class-level heterogeneity (Deb and Trivedi 2013). Alternatively, we could have estimated a Poisson fixed-effects regression using conditional likelihood. We do not because several classes have little variation in drug trials and/or failures over time, and the fixed-effects Poisson model drops these classes, sacrificing a great deal of data. In robustness checks, we show that estimates from Equation (4) are comparable to an OLS model with drug class fixed effects (a specification in which drug classes are not dropped for lack of variation in the outcome).

If β_1 is positive for both trials and failures, then we can conclude that both drug candidates and failures grow with market size. To calculate marginal failure risk, recall from our theoretical framework that $r(i^*) = \frac{\frac{\partial}{\partial m}\phi(i)}{\frac{di^*}{dm}}$, the marginal increase in failures due to market size growth, divided by the marginal increase in drug candidates. To recover

this parameter, we estimate Equation (4) for the number of phase-specific failures and the number of phase-specific drug candidates, respectively. Using the estimated parameters, we calculate the marginal change in predicted trials and failures that result from a 50% increase in ENPV in each class.[12] These changes in trials and failures then allow us to identify the marginal failure rate at any phase of development. For example, the failure rate from Phase I to II is estimated as the ratio between the marginal increase in failures from Phase I to II and the marginal increase in Phase I trials in response to an incremental increase in ENPV.

We identify the average failure rate by estimating cumulative failures and dividing by cumulative drug candidates. Using Equation (4), we predict the total number of trials and failures across classes had market size remained at pre-Part D levels (i.e., $\sum_{d\in\{c_t\}}\widehat{ENPV}_d^0$), the ratio of which forms the pre-Part D average failure rates by year. We repeat this same exercise with post-Part D market size (i.e., $\sum_{d\in\{c_t\}}\widehat{ENPV}_d^1$) to calculate the post-Part D average failure rate by year. Comparing these average failure rates will allow us to test whether the average risk of failure is rising with revenues (Implication 1). Moreover, with estimates for both the marginal and average failure rates, we can test whether the marginal failure rate exceeds the average (Implication 2) and thus whether market size drives up failure rates and launch costs. This yields insight into the slope of the pharmaceutical innovation supply curve.

Note that the covariate of interest in the second-step regression is a prediction from the first-step regression. Consequently, the second-step standard errors require adjustment. Note also that the first-step regression is estimated using data at the drug-

[12] Part D corresponded to an average 43% increase in ENPV per class.

year level while the second-step regression is estimated at the class level.[13] Therefore, standard analytic adjustments to standard errors are not feasible. Consequently, we use a non-parametric bootstrap method that uses randomly sampled (with replacement) draws of class-level clusters of drug-year observations to estimate the parameters of the first- and second-step regressions. Our coefficients of interest are estimated from the original data, and the 90%, 95%, and 99% empirical confidence intervals are estimated from 1000 bootstrap replications.

As alluded to above, we conduct several robustness checks. First, we estimate the second-step regression using an OLS fixed-effects approach. Second, we assess the impact of shortening and lengthening the duration of clinical trial development from 12 years to 10 or 14 years (i.e., if the phase had not occurred within 10 or 14 years, we count the trial has having failed).[14]

D. Predicting the Clinical Quality of Innovation

Finally, we assess how Part D affected the quality distribution of drugs. As described in Section IV, we leverage data on clinical benefits to classify new drugs in one of three categories: major, minor, and no improvements in clinical benefit. To account for the categorical nature of our quality data, we estimate a multinomial logistic regression at the drug level that compares the quality of approved drugs in classes with higher MMS relative to lower MMS, before and after Part D. Our regression takes the form

$$Q_{\mathrm{it}} = \gamma_0 + \alpha_{1\mathrm{k}}\mathrm{Part\ D_t} * \mathrm{MMS_c} + \gamma_2\mathrm{Part\ D_t} + \gamma_3 MMS_c, \qquad (5)$$

[13] While it may seem intuitive to model development outcomes at the drug level, class-level changes in revenue cannot be applied to a single drug without accounting for differences across classes in the number of new drugs being developed (Solon, Haider, and Wooldridge 2015). In larger classes, for instance, a given growth in class-level revenue will exert a weaker effect on innovation incentives than in smaller classes. One might be tempted to weight each drug by the number of drugs in its class, but this involves weighting by an endogenous variable. To avoid this conundrum, we estimate our second-stage regression at the drug-class level.

[14] Dropping the development of combination therapies does not affect our estimated results. Unfortunately, we are unable to reliably identify the development of new chemical entities.

Note that α_{1k} measures changes in the probability of observing major ($k = 1$) and minor improvement drugs ($k = 2$) relative to the probability of observing no improvement drugs.

We synthesize Part D's impact on the number of new approvals and the quality distribution of those approvals by using Equations (3) and (4) to predict the number of new approvals for each class as a result of a $1 billion increase in drug revenues and then using Equation (5) to identify the average change in quality with and without Part D. For example, we calculate the number of new drugs of major clinical benefit before Part D and then after Part D following a $1 billion increase in revenues. The difference of these two values identifies the change in the number of approved drugs with major clinical benefit as a result of a $1 billion increase in revenues. All confidence intervals are estimated from 1000 bootstrap replications with draws of class-level clusters.

IV. Data

We rely on three main sources of data for: 1) revenues; 2) innovation; and 3) drug quality.

A. Revenue Data

We use data from the Medical Expenditure Panel Survey (MEPS) from 1996 to 2017 to examine how Medicare Part D affected drug revenues across different drug classes. As our first-step estimation focuses on how Part D affected drug revenues over a 11-year life cycle among drugs introduced prior to Part D (i.e., 2006), the MEPS data frame allows us to observe all relevant years of revenues. A nationally representative survey of the US civilian non-institutionalized population, the MEPS contains information on individuals' insurance status, such as Medicare, Medicaid, and private coverage, as well as their prescription drug fills within the year. Each prescription fill has an associated

national drug code (NDC), which is the Food and Drug Administration's (FDA) unique identifier for a drug's manufacturer, ingredient, and packaging, the amount paid, and each drug is classified into a Multum therapeutic classification (i.e., "drug class"). We use the MEPS data to calculate (1) survey-weighted fractions of prescriptions filled by Medicare beneficiaries within a drug class between 2001 and 2003 (which we term the Medicare market share), and (2) survey-weighted drug-level brand revenues from 1996 to 2017 for drugs approved prior to Part D in 2006.

Since revenues vary over the life cycle of a drug and with the presence of generic competitors, we supplement the MEPS data with data from the FDA's National Drug Code (NDC) Directory, which provides information on each approved drug's FDA application type and approval dates. Branded drugs are approved via new drug applications (NDAs), whereas generics are approved via abbreviated new drug applications (ANDAs). We use this information to identify the year of generic entry for a given drug ingredient.

B. Innovation Data

Our second primary data source is PharmaProjects, documenting drug development events from January 1, 1996, to June 30, 2020. Using various public sources including international clinical trial databases, patent filings, press releases, conference proceedings, medical literature, and direct contacts with pharmaceutical companies and researchers, PharmaProjects provides a comprehensive global data source for human clinical trials. PharmaProjects reports on when drug candidates initiate Phase I, II, and III, and on when an approval occurs. While underreporting may still persist, particularly for drugs in early-stage development, we note that the bias will not affect our estimates as long as underreporting does not change systematically before and after Part D or

systematically across drugs with differential pre-Part D Medicare market share (Adams and Brantner 2006; Blume-Kohout and Sood 2013).

As noted, our dataset inevitably censors the development phase of drugs that entered the clinical trial pipeline, particularly during the later years of the sample. We assess the impact of this censoring by examining the duration of clinical development from Phase I to approval in Appendix Table B1. Duration varies across drugs, though the median duration is approximately eight years, which aligns with previously established estimates in the literature (Adams and Brantner 2006; DiMasi 2001; Wong, Siah, and Lo 2019, among others). This finding suggests that for drugs with Phase I start dates in 2012, we observe the entire clinical development process for approximately half the cohort. For this reason, we limit the latest year of analysis to drugs with Phase I start dates in 2012. Our main analysis sample relies on drugs with Phase 1 start dates between 1996 and 2012, and we follow drug development for 10, 12, and 14 years.[15] For the universe of drugs in PharmaProjects, we use the dates of "major changes in drug status" variable to identify whether and when the drug progressed from Phase I to II, Phase II to III, and Phase III to approval. If, for example, a drug does not progress from Phase I to II, we classify that as a failure, and we define the failure rate as the number of failures from Phase I to II divided by the number of trials in phase I. We apply this definition for each phase progression. Note that failing between the current phase and the next requires that the drug has progressed to at least the current phase.

PharmaProjects provides additional information on the therapeutic classification for each drug. Because PharmaProjects relies on its own proprietary therapeutic classification system, we create a crosswalk to the Multum therapeutic class system that is used in MEPS, relying on using Multum and PharmaProjects class descriptions.

[15] Prior to Part D in 2000, 14 years corresponds to the 96th percentile of duration from Phase I to approval

Appendix B: Data Details provides the crosswalk that we developed, as well as additional information about the data.[16]

C. Clinical Quality Data

Finally, we follow the prior literature and identify the quality of a drug using France's Haute Autorité de Santé's (HAS) measures of clinical benefit (Krieger, Li, and Papanikolaou 2021; Kyle and Williams 2017). Measuring clinical benefit is challenging across a range of drugs, because diseases differ in their clinical endpoints. One approach is to rely on published cost-effectiveness analyses that measure quality-adjusted life-years gained. However, cost-effectiveness studies in the published literature vary widely in their methodological quality, their methodological approach (e.g., the components of value considered), and in the approach to choosing a comparator against which to compute incremental gains in health. HAS quality evaluations eliminate many of these biases, because all drugs in France need to be reviewed by the HAS to be included on a list of reimbursable products, and a uniform assessment process is used.

We rely on HAS assessment of a drug's clinical value relative to the best existing therapy, known as the amélioration du service médical rendu (ASMR) or the improvements in medical benefit. HAS ASMR evaluations—which are based on data from the clinical trial, meta-analyses, observational studies, and filings from the pharmaceutical companies—are estimated on a five-point scale which identifies whether a drug offers major (ASMR = 1), significant (ASMR = 2), moderate (ASMR = 3), minor (ASMR = 4), or no improvements (ASMR = 5). The highest category of value is reserved for drugs with a demonstrated impact on mortality. Lower categories of value are then

[16] Our final PharmaProjects data consist of a balanced panel of drug developments and failures by class-year. In our sensitivity analyses, we identified firms with Medicare revenues prior to Part D by matching firm names from MEPS to PharmaProjects. This crosswalk, along with other details, are also included in Appendix B: Data Details.

assigned based on a drug's additional clinical effect measured in terms of efficacy and tolerance.

Of note, drugs in ASMR categories 1 through 3 (i.e., major, significant, and moderate improvements) receive expedited market access and favorable pricing. Moving forward, we simply refer to drugs in these three categories as generating "major" clinical improvement. In contrast, drugs with minor clinical benefit (i.e., ASMR category 4) can only be priced above its comparator in limited circumstances.[17] Drugs in the lowest category of clinical benefit (i.e., "no" improvement) are reimbursed only if their costs are lower than their comparator's (Haute Autorité de Santé 2014).

For some drugs, we observe multiple HAS evaluations. Outside of the first HAS evaluation, subsequent HAS evaluations occur when a new indication has been added or when the original indication needs to be re-evaluated: HAS evaluations are required every 5 years to maintain reimbursement. When there were multiple HAS evaluations, we use the drug's highest observed measure of clinical benefit (i.e., the lowest ASMR category). We identified each drug's first date of clinical development by matching drugs in HAS to those in the PharmaProjects dataset.

V. Results

A. Summary Statistics

Table 1 shows the summary statistics for the MEPS (panel A), PharmaProjects (panel B), and HAS data. Panel A shows variables from the MEPS dataset at the drug-year level, with the first two columns averaging over 1996 to 2005 and the last two columns representing averages between 2006 to 2017. The average revenue pre-Part D was

[17] A drug can be priced higher than its competitor if it has a different target population.

about $207 million per year. After part D, these drugs approved which had been approved before 2006 had an average revenue of $497 million per year. The average Medicare market share was almost 40% and about 43% of drugs had a generic competitor. The average approval year ranged from 1990 to 1998, and drugs were on average 10 to 12 years post approval. Panel B shows variables from the PharmaProjects dataset at the class-year level, with the first two columns averaging over 2001 to 2005 and the last two columns averaging 2006 to 2012. Across each of the 96 drug classes, there were on average 1.3 to 2.1 Phase I trials, 0.6 to 0.9 Phase II trials, 0.2 to 0.4 Phase III trials, and 0.1 to 0.2 drug approvals per class-year. Finally, panel C shows that prior to Part D, 20% of drugs offered major improvements, 20% contributed only minor improvements, and 60% had no improvement in clinical value. After Part D, the probability of drugs making major and minor improvements increased to 23% and 25%, respectively, while drugs with no improvement fell to 52%.

B. Effect of Part D on Revenue Growth

[Insert Table 2c Here]

To estimate the causal impact of Part D on R&D investments and failure rates, we begin by identifying how Part D affected revenues. Table 2 shows the estimates of Equation (3). Both the enactment of Part D in 2003 and its implementation in 2006 increased drug-level revenues, though the anticipatory response is close to zero when drug-, year-, and approval-year fixed effects are included in the model. Importantly, we find that Part D increased revenues more for drugs with higher Medicare market share: 10 percentage points of additional Medicare market share generates a 15.5% differential increase in post-Part D revenues, corresponding to a marginal effect of $36.0 million per

drug.[18] The test statistic for the hypothesis that the coefficient on the interaction of the Part D indicator and MMS is zero is rejected with a χ^2 value of 5.14 that corresponds to a p-value of 0.023.[19] As expected, a generic entrant in the market reduces revenues, which falls by 32%.

[Insert Figure 5 Here]

Using the estimates shown in Table 2, we calculate predicted revenues with and without Part D for each drug within the class. We sum across drugs to get class-level revenue over time, before and after Part D. Figure 5 plots the positive relationship between Medicare market share and the implied percentage change in class-level revenue due to Part D implementation. The average drug class had a Medicare market share of 42% and a corresponding 146% increase in revenues due to Part D. Drug classes with the highest Medicare market share include cholinesterase inhibitors (used in Alzheimer's treatment), coagulation modifiers, antiarrhythmic agents, inotropic agents (for treating heart failure), chelating agents, and antianginal agents. Among this group of drugs that have Medicare market share greater than 75%, post-Part D revenues increased on average 224%. In contrast, drug classes with the lowest Medicare market share include central nervous system stimulants (used to treat disorders such as attention deficit hyperactivity disorder), anorexiants, antituberculosis agents, tetracyclines, and penicillins. Among this latter group, which had Medicare market shares less than 10%, post-Part D revenues increased, on average, 15.6%.

---

[18] Increasing MMS by 1 (i.e., 100 percentage points) yields an increase of post-Part D revenues by exp(0.936)-1 = 165%.

[19] Although the model in the second step of our analysis should not be construed as an instrumental variables regression, it is comforting to note that this χ^2 value suggests a strong "instrument". In fact, when we estimate the regression in the first step using OLS on the logarithm of revenue, the "first stage" F-statistic equals 31.35.

C. Effect of Revenue Growth on R&D Investments and Failure Rates

[Insert Table 3 Here]

With estimates from the first step in hand, we turn to estimates from our model in the second step of our analysis, as specified in Equation (4). Table 3 presents the elasticities of clinical trials (panel A) and failures (panel B) with respect to ENPV. For a 1% increase in ENPV, the number of Phase I, II, and III trials (within 12 years of the Phase I start date) increase by 0.14%, 0.20%, and 0.23%, respectively. All these estimates are all statistically significant at the 5% level; the Phase I and II estimates are also significant at the 1% level. For comparison, Dubois et al (2015) estimate that the elasticity of drug approvals with respect to global ENPV at the class level is 0.23. While their measurement concept differs slightly from ours, the magnitudes are similar. We obtain 0.26, and their estimate lies within our confidence interval. Medicare Part D increased ENPV by 43%, so the elasticities we estimate imply that Part D increased annual drug approvals by 11.2%.

We observe a similar positive elasticity of the number of clinical trial failures with respect to market size. For a 1% increase in ENPV, the number of failures from Phases I to II, Phases II to III, and Phases III to approval increase by 0.19%, 0.23%, and 0.23%, respectively. Again, estimates are statistically significant, particularly at earlier stages of clinical trial development. The failure elasticity from Phase I to approval is 0.14. Thus, in response to a 43% Part-D induced increase in ENPV, failures from Phase I to approval also increased by 6.3% per year. Holding the number of Phase I trials constant, this means that Part D increased the number of failures from 140 of the 152 Phase I trials to almost 149 of the 152 Phase I trials. However, as we show above, the number of Phase I trials are also increasing with market size.[20]

[20] OLS estimates also demonstrate a positive relationship between ENPV and trials and failures. Shown in

[Insert Table 4 Here]

To understand whether failures are increasing faster than trials, we calculate, in Table 4, the marginal and average failure rates in response to a change in market size. As described in Section II.B.1, the marginal failure rates are calculated by dividing marginal changes in trials and failures as a result of a 50% increase in ENPV, whereas average failure rates are calculated by dividing total predicted trials and failures with a 1% increase in ENPV. The marginal drug induced by a 50% increase in market size exhibits a 79% likelihood of failing between Phases I and II. In contrast, the average drug has a much lower failure rate at 58%. The 95% confidence intervals of the marginal and average failure rates do not overlap, suggesting that the marginal failure rate is statistically higher than the average failure rate and that average failure rates are indeed rising with market size.[21] Conditional on making it to Phase II, the marginal drug has an 87% likelihood of failure, while the average drug has a 77% chance of failing again consistent with upward-sloping supply. The confidence intervals of these two estimates overlap only minimally. The one exception occurs for drugs that progress to Phase III, where the marginal and average drug failure rates were statistically indistinguishable, with overlapping confidence intervals.

In column (4) of Table 4, we show the overall failure rate from Phase I to approval: the marginal drug entering Phase I has a 98% chance of failing, which is statistically higher than the average drug with an overall failure rate of 92%. While this

---

Appendix Table C1, a 10% increase in ENPV results in the number of Phase I, II, and III trials increasing by 1.24, 0.45, and 0.20, respectively. Failures from Phase I to II, II to III, and III to approval also increase by 0.81, 0.37, and 0.12. The number of failures from Phase I to approval increase by 1.15. Elasticities from the OLS estimates are sensitive to the pre-period mean, which differs significantly across classes. As such, we directly report the coefficients from the OLS fixed-effect specification.

[21] Note that comparing marginal failure rates to contemporaneous average failure rates is more stringent than comparing marginal failure rates post-Part D with pre-Part D average failure rates.

difference may appear small at first glance, recall our earlier point that small changes in failure rates translate into large changes in the number of trials needed to develop one drug.[22] Based on these estimated failure rates, the average drug needs an average of 12 Phase I trials to produce one approval. In contrast, the marginal drug needs more than 45 Phase I trials to produce an average of one approval.

[Insert Figure 6 Here]

We provide additional clarity on the differences in the marginal post-Part D and average pre-Part D failure rates in Figure 6. The figure confirms that the marginal failure rates post market size expansion are significantly different from the average failure rates prior to the market size increase (i.e., the addition of marginal drugs increase the average failure rate). The differences between marginal post- and average pre-expansion failure rates are large and statistically significant, particularly from Phases I to II, Phases II to III, and Phases I to approval. Note, however, that the average failure rate does not change meaningfully in the post-period. We note that with larger increases in ENPV, the arrival of marginal drugs will increase post-period average failure rates by even more.

Our findings support the theoretical implications discussed in Section II. In particular, we have demonstrated that the exogenous increase in market size from Part D led to increases in the average risk of failures and NND (Implication 1). We have also shown that the post-Part D marginal failure risk exceed even the post-Part D average failure risk (Implication 2), which implies the failure risk is increasing monotonically with each R&D investment, and that – if trial costs do not decrease with market size – the supply of pharmaceutical innovation is rising locally.

[22] From the OLS approach, we divide the failures coefficients by the trials coefficients to identify the marginal failure rates, which results in a 93% failure rate from Phase I to approval for the marginal drug candidate (Appendix Table C1).

Trials in our main specification defined as having failed if they did not proceed within 12-years of the Phase I start date. To assess whether our estimates are sensitive to this censoring, we additionally consider the trails, failures, and approvals within 10- and 14-years of the phase I start. In Appendix Table C2, we find comparable estimates to our main results. Elasticities of approvals with respect to ENPV range from 0.75 to 0.18, and elasticities of failures from Phase I to approval remain at 0.18.

D. Implications of demand growth for R&D costs and innovation quality

[Insert Table 5 Here]

In the final section of empirical results, we shed light on how launch costs have changed with increases in market size as well as the quality of new drugs launched. We first identify the number of Phase I, II, and III trials required to develop the marginal approved drug and the average pre-part D drug. The number of Phase X trials needed for the marginal approved drug equals the inverse of the probability of succeeding from Phase X to approval, where these probabilities come from Table 4. We report the number of trials needed in each phase in Table 5 Next, we use estimates from the literature—including DiMasi, Hansen, and Grabowski (2003),Adams and Brantner (2006), DiMasi, Grabowski, and Hansen (2016) and (Sertkaya et al. 2024)—to identify the mean "out-of-pocket" costs borne by innovators for each phase of development. With these numbers, we can then identify the costs to launch a new drug that are shown in Table 5.

Prior to Part D, launching one drug required on average 29 phase I trials, 13 phase II trials, and 3 phase III trials. However, post-Part D, successfully launching one marginal drug required 106 phase I trials, 22 phase II trials, and 3 phase III trials. The significant increases in phase I and II trials needed to launch one drug drives the increase in the marginal cost of innovation. The magnitude of this cost increase depends

on the literature estimates applied. For example, using clinical trial cost estimates from (Adams and Brantner 2006), we find that launching the marginal drug will cost $0.4 billion more than launching the average drug: a difference between $1.3 billion (marginal) and $0.87 billion (average). On the other hand, using estimates from (DiMasi, Grabowski, and Hansen 2016) leads to a $2.6 billion increase in launch costs, where the marginal drug costs $5.1 billion and the average drug costs $2.4 billion.

[Insert Table 6 Here]

Next, we turn to our estimates of Equation (5) to assess changes in the quality distribution of approved drugs. Shown in Table 6, Part D increased the likelihood that drugs explored had major improvements in clinical value by 11.6 percentage points, whereas drugs with no clinical improvement fell by 14.4 percentage points. The likelihood of a drug with minor improvement increased by a statistically insignificant 2.8 percentage points. Applying these changes in probability to the number of new drugs introduced as a result of a $1 billion increase in ENPV, we find that we will have 1.6 more drugs with major benefit and 0.9 fewer drugs with no clinical benefit. There will be a statistically insignificant increase of 0.5 more drugs with minor clinical benefit, resulting in an overall increase of 1.2 new drugs launched.

E. Implications for Pharmaceutical Industry Rents

1. Empirical Framework

Upward-sloping supply of innovation implies that the average drug costs less to launch than the marginal drug. This results in Ricardian rents, i.e., rents generated by differences in productivity, rather than by market power itself (Gochberg and Menaldo 2022; Shaanan 2006). In this context, differences in productivity correspond to differences in the probability of success and in the post-launch profits of an innovation. Our

theoretical framework allows us to characterize these Ricardian rents, and, with a bit more structure than we have needed so far, recover a lower bound on them.

Notice that expected Ricardian rents from drug candidate i are given by $E(Rent(i)) \equiv \sigma(i)\pi(i;m) - \delta$. Aggregate expected rents for the industry are given by:

$$Rent(m) \equiv \int_0^{i^*(m)} \sigma(s)\pi(s;m)\, ds - \delta i^*$$

δi^* equals total launch cost. Therefore, recalling that $\ell(m)$ yields the number of drugs launched when market size is m, we can express aggregate Ricardian rents as:

$$Rent(m) \equiv \int_0^{i^*(m)} \sigma(s)\pi(s;m)\, ds - \ell(m)ALC(i^*)$$

So far, we have allowed notional demand, $\pi(i;m)$, to slope downward or upward. If instead we restrict it to slope downward only, a lower bound on $Rent(m)$ can be estimated. Specifically, assume that:

$$\pi_i(i;m) \leq 0 \qquad \text{(Condition 3)}$$

Profits depend on a host of factors, including clinical quality, the willingness to pay for clinical improvements, and the size of the disease indication being treated. Each of these factors may vary with i. We have provided evidence that clinical quality rises with i, which would bend the slope of the demand curve upward, but we cannot rule out the possibility that willingness to pay per unit of health improvement,[23] indication size, or other similar factors fall with i. Thus, it remains *a priori* unclear whether $\pi(i;m)$ slopes upward or downward in i.

If Condition (**Error! Reference source not found.**) holds, then this combined with Condition (1) implies:

$$Rent(m) > \sigma\big(i^*(m)\big)[\pi(i^*(m);m)i^*(m) - ALC(i^*(m))i^*(m)] \qquad (6)$$

[23] For instance, if there are diminishing or increasing returns to health improvement, the willingness to pay for a unit of health improvement might vary with disease severity (Lakdawalla and Phelps 2020).

Since $\pi(i^*(m); m) = MLC(i^*)$ and $\ell(m) = \sigma\big(i^*(m)\big)i^*(m)$, this expression can be equivalently rewritten as:

$$Rent(m) > \ell(m)[MLC(i^*(m)) - ALC(i^*(m))] \quad (7)$$

With downward-sloping demand, the post-launch profits that accrue from each launched drug exceed $MLC(i^*)$, but the average cost of a launched drug is $ALC(i^*)$. Therefore, the rents on the average drug candidate will be at least $[MLC(i^*(m)) - ALC(i^*(m))]$, and total rents will be $\ell(m)[MLC(i^*(m)) - ALC(i^*(m))]$. This yields an empirically estimable lower bound on the Ricardian rents accruing from upward-sloping supply, also depicted as the area of the shaded "Min. Rents" box in Figure 7.

2. Empirical Measurement of Ricardian Rents

To measure Ricardian rents, we calculate the marginal R&D cost share, defined as the marginal cost of launching new drugs as a share of the change in ENPV. For the denominator, we note that the average total ENPV per class pre-Part D was $10.9 billion, and a 50% increase in ENPV equates to a $5.45 billion increase in payments. For the numerator, we note that the marginal cost of launching one drug costs $5.1 billion (Table 5 using DiMasi, Grabowski, and Hansen (2016)). Discounting this cost over 10 years using a discount rate of 10% leads to a capitalized launch cost for one drug of $12 billion.[24] However, the average increase in annual approvals per class was only 0.014 drugs. This makes the annual increase in capitalized launch costs equal to $0.168 billion, and the long-run increase in capitalized launch costs using a 10% discount rate yields $1.85 billion.[25] Therefore, the marginal launch cost as a fraction of ENPV is 33.9% ($1.85

[24] Because the first year is treated as year 0, this leads to \$5.1 billion*$(1.1)^9$ = \$11.989 billion.

[25] We arrive at the long-run increase in capitalized launch costs with (\$11.989 billion * 0.014)/(1.1/0.1) = \$1.85 billion.

billion / $5.45 billion). Applying this marginal R&D cost share to the average launch cost of $2.5 billion, we then arrive at innovator rents as a fraction of ENPV at 15.6%.[26]

Our estimate for the marginal R&D cost share is broadly consistent with literature estimates of pharmaceutical production, distribution, and marketing costs. Sood et al. (2017) find that pharmaceutical distribution expenses account for 24% of revenues. Lakdawalla, Philipson, and Wang (2007) estimate that marketing accounts for 19% of revenues, and the US Government Accountability Office (2012) has reported that production costs account for 25% of revenues. This leaves a gross margin of 32%. In equilibrium, we would expect that the marginal R&D cost is equal to downstream revenues net of all variable costs, and our estimate of 33.9% is aligned with these back-of-the-envelope calculations based on the existing literature.

VI. Conclusions

In this paper, we show that growth in market size increases R&D failure rates and costs, but with an accompanying increase in clinical quality. We interpret these effects as the result of upward-sloping innovation supply interacting with growth in the financial incentives to innovate. Under the assumption that the marginal drug candidate is less profitable than its predecessors, we can also derive a lower bound on Ricardian rents in the pharmaceutical industry. This analysis implies that productivity differences, rather than barriers to entry, account for half of pharmaceutical profits or approximately the entire "excess return" observed in the pharmaceutical industry.

From our analysis, we conclude that the cost of launching the marginal drug post part-D is around $2.76 billion, and the elasticity of new approvals relative to market

[26] This is calculated with (0.339*$2.5 billion)/$5.45 billion.

size is 0.18. This estimate is in line with earlier studies. Using global data from 1997 to 2007, Dubois et al. (2015) find that $2.5 billion in revenues stimulates one new chemical entity, and the elasticity of new chemical entities relative to revenue is 0.23. But importantly, beyond the revenues needed to develop one drug, we uncover interesting facts that inherently imply increased costs of development. The marginal drugs induced by Part D-related market size expansion failed at a 98% rate, significantly higher than the failure rate for the average pre-Part D drug. Consequently, the average failure rate during the post-Part D period increased steadily. The marginal post-Part D drug experienced increases in failure risk, particularly at the later stages of development. The growth in failure coincided with increases in quality, suggesting that there were gains to taking the additional risk.

Our estimates imply that the marginal drug costs 40% more to launch than the average drug, suggesting a substantial upward slope to the pharmaceutical innovation supply curve. As market size increases, the launch cost of the marginal drug increases, the average launch cost is driven up, and the rents captured by the industry increase. In these calculations of launch cost, we have held constant the cost of drug developments in each phase over time, and we ignore the opportunity cost of time. Accurately measuring the cost of time requires information on whether Part D accelerated phase lengths or the mean length of time between phases. If so, total costs post-Part D may be smaller than actual cash outlays. But to the extent that costs are increasing over time due to increased clinical trial complexity, larger clinical trial sizes, and possibly changes in protocol design, our launch cost estimates will understate true costs (Getz and Kaitin 2015; Getz et al. 2008; Viereck and Boudes 2011).

We contribute a novel explanation for excess profits in the pharmaceutical industry. Upward-sloping supply results in profits for inframarginal drugs that are not easily

competed away, even when considering the entry of new drugs. Inframarginal profits are also difficult for policymakers to reduce without collateral effects on innovation rates. Reductions in demand through price caps or other related policies will make the marginal drug unprofitable, resulting in lower innovation even though such policies may simultaneously wring some profits out of inframarginal drug candidates.

Our study points towards several future research questions. What is the value of the marginal drug to consumers, and how does this compare to its marginal launch cost? This sheds light on the consumer surplus generated by innovation on the margin. What is the shape of the notional demand curve, and specifically, how does profitability change on the margin for new innovations? This requires an experiment that shifts innovation supply, rather than demand. Finally, are pharmaceutical rewards too high, too low, or approximately efficient? Put differently, at what point do price increases stimulate innovation whose social cost approximately equals its social value? This final question ranks as one of the most important outstanding questions in the economics of the pharmaceutical industry (Lakdawalla 2018).

VII. References

Acemoglu, Daron, and Joshua Linn. 2004. 'Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry', *Quarterly Journal of Economics*, 119: 1049-90.

Adams, Christopher P., and Van V. Brantner. 2006. 'Estimating the Cost of New Drug Development: Is it really $802 Million?', *Health Affairs*, 25: 420-28.

Angell, Marcia. 2004. *The Truth About the Drug Companies: How They Deceive Us and What to Do About It* (Random House: New York).

Aronson, J. K., and A. R. Green. 2020. 'Me-too pharmaceutical products: History, definitions, examples, and relevance to drug shortages and essential medicines lists', *Br J Clin Pharmacol*, 86: 2114-22.

Bennette, Caroline Savage, Anirban Basu, 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: 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.

Deb, Partha, Edward C Norton, and Willard G Manning. 2017. *Health Econometrics Using Stata* (Stata Press: College Station, Texas).

Deb, Partha, and Pravin K Trivedi. 2013. 'Finite mixture for panels with fixed effects', *Journal of Econometric Methods*, 2: 35-51.

DiMasi, Joseph A, Henry G Grabowski, and Ronald W Hansen. 2016. 'Innovation in the pharmaceutical industry: new estimates of R&D costs', *J Health Econ*, 47: 20-33.

DiMasi, Joseph A, Ronald W Hansen, and Henry G Grabowski. 2003. 'The Price of Innovation: New Estimates of Drug Development Costs', *J Health Econ*, 22: 151-85.

DiMasi, Joseph A. 2001. 'New drug development in the United States from 1963 to 1999', *Clinical Pharmacology & Therapeutics*, 69: 286-96.

Dranove, David, Craig Garthwaite, and Manuel I Hermosilla. 2020. "Expected profits and the scientific novelty of innovation." In.: National Bureau of Economic Research.

Dubois, Pierre, Olivier de Mouzon, Fiona Scott-Morton, and Paul Seabright. 2015. 'Market size and pharmaceutical innovation', *RAND Journal of Economics*, 46: 844-71.

Duggan, Mark, and Fiona Scott Morton. 2010. 'The Effect of Medicare Part D on Pharmaceutical Prices and Utilization', *American Economic Review*, 100: 590-607.

Finkelstein, Amy. 2004. 'Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry', *Quarterly Journal of Economics*, 119: 527-64.

Getz, Kenneth A., and Kenneth I. Kaitin. 2015. 'Chapter 1 - Why Is the Pharmaceutical and Biotechnology Industry Struggling?' in Peter Schüler and Brendan Buckley (eds.), *Re-Engineering Clinical Trials* (Academic Press: Boston).

Getz, Kenneth A., Julia Wenger, Rafael A. Campo, Edward S. Seguine, and Kenneth I. Kaitin. 2008. 'Assessing the Impact of Protocol Design Changes on Clinical Trial Performance', *American Journal of Therapeutics*, 15.

Gochberg, Will, and Victor Menaldo. 2022. 'To rent or not to rent? Mechanics, causes and consequences of Ricardian and Quasi-rents in the oil industry', *Resources Policy*, 78: 102826.

Haute Autorité de Santé. 2014. 'Pricing and Reimbursement of drugs and HTA policies in France'. https://www.has-sante.fr/upload/docs/application/pdf/2014-03/pricing_reimbursement_of_drugs_and_hta_policies_in_france.pdf.

Hell, Steven C, Samuel H Zuvekas, and Marc W Zodet. 2011. 'Implications of the Accuracy of MEPS Prescription Drug Data for Health Services Research', *Inquiry*, 48: 242-59.

Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou. 2021. 'Missing Novelty in Drug Development*', *The Review of Financial Studies*, 35: 636-79.

Kyle, Margaret, and Heidi Williams. 2017. 'Is american health care uniquely inefficient? evidence from prescription drugs', *American Economic Review*, 107: 486-90.

Lakdawalla, Darius N. 2018. 'Economics of the pharmaceutical industry', *Journal of Economic Literature*, 56: 397-449.

Lakdawalla, Darius N., and Charles E. Phelps. 2020. 'Health Technology Assessment with Risk Aversion in Health', *J Health Econ*, 72.

Lakdawalla, Darius; , Tomas; Philipson, and Y. Richard Wang. 2007. 'Intellectual Property and Marketing', *NBER Working Paper*, No 12577.

Ledley, Fred D, Sarah Shonka McCoy, Gregory Vaughan, and Ekaterina Galkina Cleary. 2020. 'Profitability of large pharmaceutical companies compared with other large public companies', *JAMA*, 323: 834-43.

Morgan, Daniel C., and Samuel E. Allison. 1964. 'The Kefauver Drug Hearings in Perspective', *The Southwestern Social Science Quarterly*, 45: 59-68.

Myers, Kyle, and Mark Pauly. 2019. 'Endogenous productivity of demand-induced R&D: evidence from pharmaceuticals', *The RAND Journal of Economics*, 50: 591-614.

Nordhaus, William D. 1969. 'An Economic Theory of Technological Change', *American Economic Review*, 59: 18-28.

Phelps, Charles E, and Chris Cinatl. 2021. 'Estimating optimal willingness to pay thresholds for cost-effectiveness analysis: A generalized method', *Health economics*, https://doi.org/10.1002/hec.4268.

Philipson, Tomas J, and Troy Durie. 2021. 'Issue brief: the impact of HR 5376 on biopharmaceutical innovation and patient health', *The University of Chicago*.

Philipson, Tomas J, and Anupam B Jena. 2006. 'Who benefits from new medical technologies? Estimates of consumer and producer surpluses for HIV/AIDS drugs', *Forum for health economics & policy*, 9.

Sachs, Rachel, Loren Adler, and Richard Frank. 2023. 'A holistic view of innovation incentives and pharmaceutical policy reform', *Health Affairs Scholar*, 1.

Sertkaya, Aylin, Trinidad Beleche, Amber Jessup, and Benjamin D. Sommers. 2024. 'Costs of Drug Development and Research and Development Intensity in the US, 2000-2018', *JAMA Network Open*, 7: e2415445-e45.

Shaanan, Joseph. 2006. 'Ricardian or monopoly rents? The perspective of potential entrants', *Eastern Economic Journal*, 32: 19-30.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. 'What Are We Weighting For?', *The Journal of Human Resources*, 50: 301-16.

Sood, Neeraj; , Tiffany; Shih, Karen; Van Nuys, and Dana P. Goldman. 2017. 'Follow The Money: The Flow Of Funds In The Pharmaceutical Distribution System', *Health Affairs Forefront*

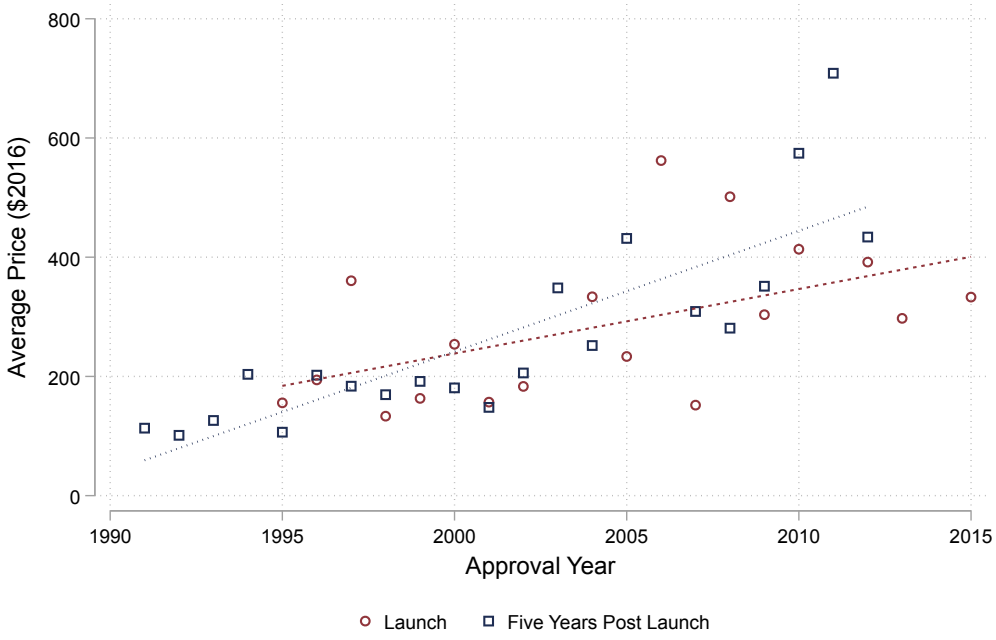
US Government Accountability Office. 2012. 'Drug Pricing: Research on Savings from Generic Drug Use'. https://www.gao.gov/assets/files.gao.gov/assets/gao-12-371r.pdf.

Viereck, Christopher, and Pol Boudes. 2011. 'An analysis of the impact of FDA's guidelines for addressing cardiovascular risk of drugs for type 2 diabetes on clinical development', *Contemporary Clinical Trials*, 32: 324-32.

Wong, Chi Heem, Kien Wei Siah, and Andrew W. Lo. 2019. 'Estimation of clinical trial success rates and related parameters', *Biostatistics (Oxford, England)*, 20: 273-86.

VIII. Tables and Figures

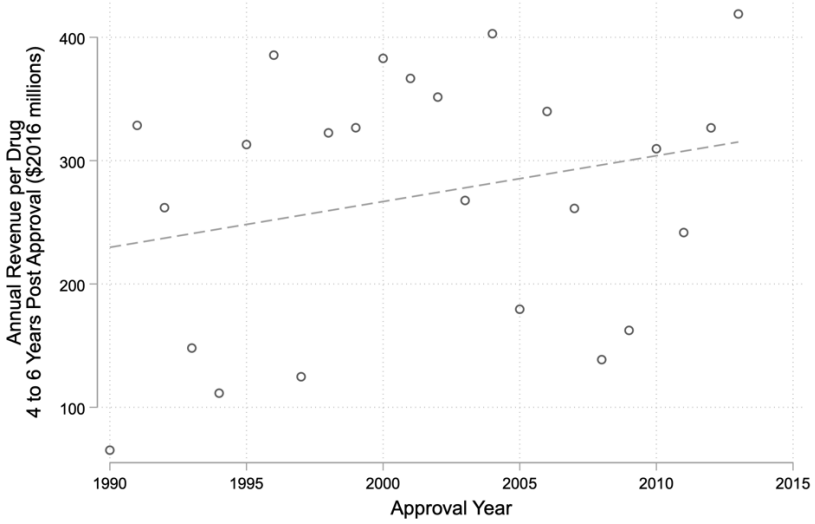
Figure 1. Average Prices at Launch and Five Years Post Launch, By Approval Year



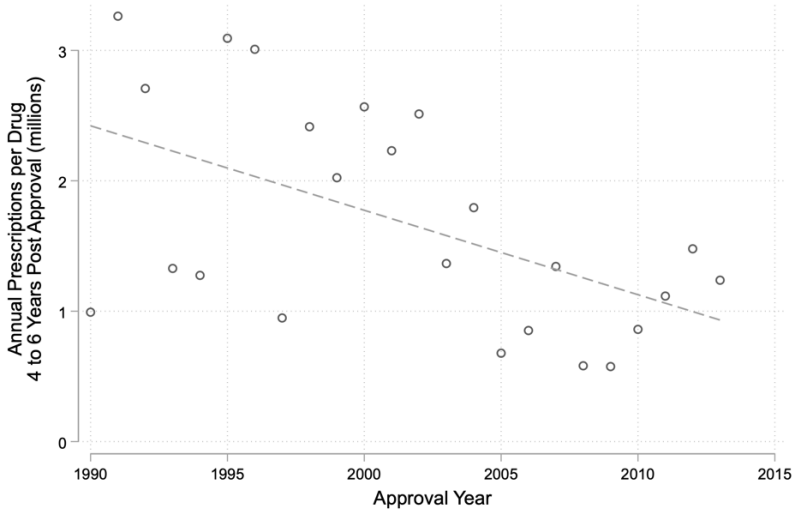
Notes: Data at the drug-year level from MEPS 1996 to 2017 and the National Drug Code Directory. The average price per drug at launch or post-launch are weighted by sales at launch or post launch, respectively. Launch prices grow, on average, at $10.82 a year (p-value 0.026) from a mean price of $284.58. Prices five years post launch grow, on average, at $20.25 a year (p-value<0.001) from a mean of $267.71.

Figure 2. Revenues and Quantities of Drugs 4 to 6 Years Post Approval

a. Annual Revenues per Drug

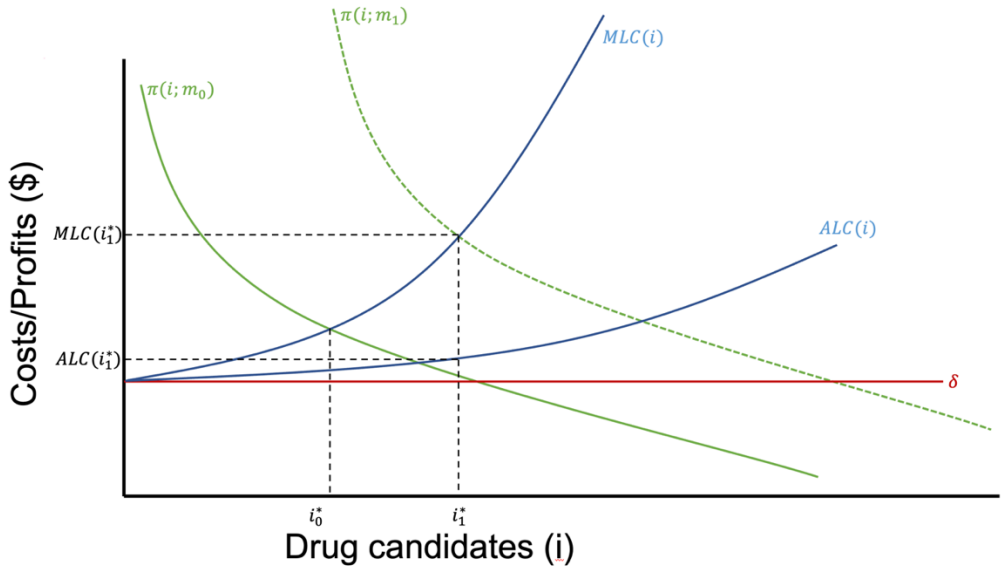


b. Annual Prescriptions per Drug



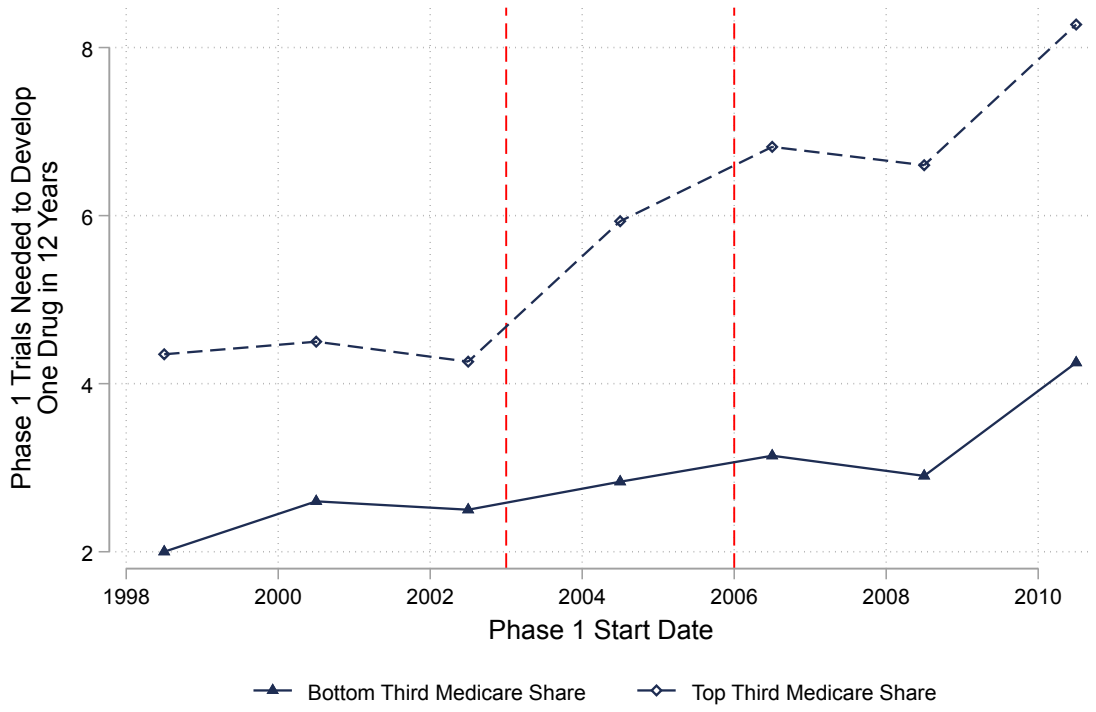
Notes: Data at the drug-year level from MEPS 1996 to 2017 and the National Drug Code Directory. Plot (a) illustrates that inflation-adjusted revenues 4-to-6 years post launch increase weakly at $3.7 million per year (p-value of 0.23) from a mean of $272 million per year. Plot (b) shows that prescriptions per drug 4-to-6 years post launch fall at a rate of 64,757 prescriptions per year (p-value of 0.007) from a mean of 1.67 million per year.

Figure 3: Effect of Demand Growth on Development Costs in the Presence of Upward-sloping Innovation Supply.



Notes: The green curves depict $\boldsymbol{\pi(i;m)}$, the post-approval profits of a type $\boldsymbol{i}$ drug when market size is $\boldsymbol{m}$. The blue curves depict the marginal cost of launching a type $\boldsymbol{i}$ drug ($\boldsymbol{MLC(i)}$) and the average cost of launching $\boldsymbol{i}$ drugs ($\boldsymbol{ALC(i)}$). $\boldsymbol{\delta}$ is the cost of researching a drug candidate.

Figure 4. Number of Phase I Trials Needed to Approve One Drug in 12 Years



Notes: Data from PharmaProjects. We show the number of trials needed to develop one drug in tewelve years over time. This is calculated as the inverse of the success rate from phase I to approval.

Table 1. Summary Statistics

	Before 2005		After 2006	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
	A. MEPS			
Revenue ($millions)	230.86	598.95	502.02	1179.25
Medicare market share (MMS)	0.39	0.18	0.40	0.18
1(Part D Passed)	0.30	0.46	1	0
1(Part D Implemented)	0	0	1	0
1(Generic Entry)	0.27	0.45	0.35	0.48
Years Since Launch	9.35	5.72	11.69	4.48
Approval Year	1991.26	6.62	1998.29	4.14
Number of Observations	3,714		2,951	
	B. PharmaProjects			
Phase I	1.267	2.626	1.979	4.817
Phase II	0.557	1.454	0.828	2.312
Phase III	0.191	0.578	0.337	0.939
Approved	0.108	0.421	0.154	0.509
Failures Phase I to II	0.710	1.429	1.161	2.816
Failures Phase II to III	0.446	1.160	0.619	1.718
Failures Phase III to Approval	0.139	0.473	0.218	0.661
Failures Phase I to Approval	1.158	2.423	1.824	4.480
Number of Observations	424		564	
	C. HAS			
Major Improvement (ASMR 1-3) (%)	19.6	39.7	23.3	42.4
Minor Improvement (ASMR 4) (%)	20.4	40.4	25.0	43.5
No Improvement (ASMR 5) (%)	59.9	49.1	51.7	50.2
Number of Observations	367		120	

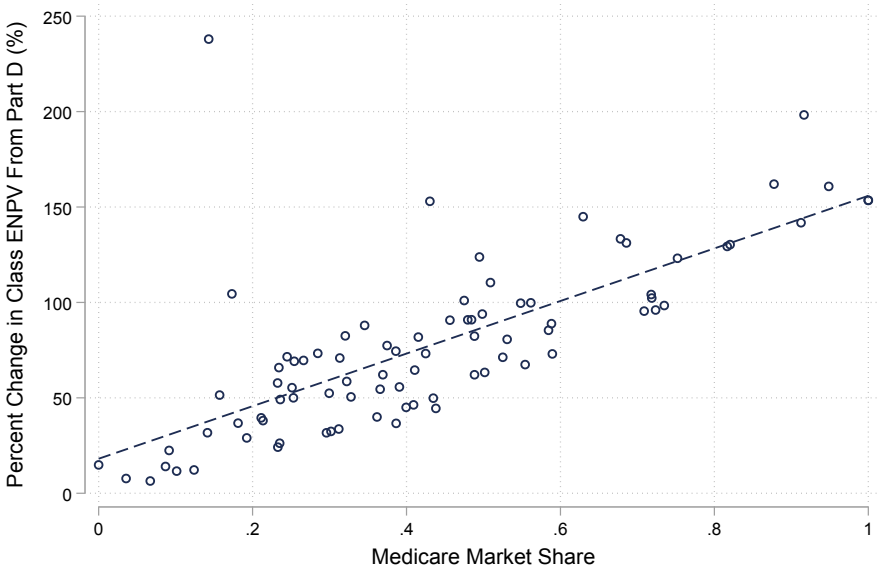
Notes: Data for the first step is at the drug-year level from MEPS and FDA's NDC Directory. For panel A, the first two columns correspond to 1996 to 2005, and the last two columns correspond to 2006 to 2017. Medicare market share is calculated using data from 2001 to 2003. MEPS sample includes drugs approved prior to 2006. Data for the second step is at the class-year level from the PharmaProjects data. For panel B, the first two columns show statistics for 2001 to 2005, whereas the last two columns show statistics for 2006 to 2012. If there has been no clinical development after a 12-year period, we treat the drug candidate as failing to progress to the next phase. Panel C shows data from HAS (matched to PharmaProjects) at the drug level.

Table 2. Impact of Medicare Part D on Revenues: First Stage Regression

	Revenues (1)	Revenues (2)	Revenues (3)
1(Part D)*MMS	1.040**	0.983**	0.936**
	(0.132 - 1.948)	(0.111 - 1.856)	(0.127 - 1.745)
1(Part D)	-0.0787	0.389	0.495
	(-0.518 - 0.360)	(-0.280 - 1.058)	(-0.229 - 1.220)
1(Part D Enacted)*MMS	-0.623	-0.632	-0.826
	(-1.907 - 0.661)	(-1.957 - 0.694)	(-2.039 - 0.386)
1(Part D Enacted)	0.643**	0.751***	0.723***
	(0.138 - 1.149)	(0.229 - 1.272)	(0.231 - 1.215)
1(Generic Entry)	-0.380***	-0.397***	-0.322**
	(-0.647 - -0.113)	(-0.662 - -0.131)	(-0.591 - -0.0534)
MMS	5.018***	5.012***	5.497***
	(3.904 - 6.132)	(3.842 - 6.182)	(4.456 - 6.538)
Drug Class FE	X	X	X
Year FE		X	X
Approval Year FE			X
Observations	6,655	6,655	6,655

Notes: Data at the drug-year level from the MEPS. Generalized linear model with a Gamma distribution and log link is estimated. Robust 95% confidence intervals in parentheses *** p<0.01, ** p<0.05, * p<0.1

Figure 5. Part-D Induced Change in Revenue by Medicare Market Share



Notes: Each dot represents a therapeutic drug class. The Medicare market share is calculated using MEPS data between 2001 and 2003. The percent change in class revenues compares predicted revenues pre-Part D between 2001 and 2003 with predicted revenues post-Part D between 2006 and 2012. The lowest Medicare market share classes include central nervous system stimulants, tetracyclines, anorexiants, and penicillins. The highest Medicare market share classes include cholinterase inhibitors, antineoplastics, and psoralens.

Table 3. Impact of ENPV on Clinical Trial Development and Failures

	Elasticity of Trials With Respect to ENPV			
	Phase I Trials	Phase II Trials	Phase III Trials	Approvals
	(1)	(2)	(3)	(4)
Log(ENPV)	0.136***	0.198***	0.232**	0.266**
	[0.06, 0.17]	[0.06, 0.25]	[0.04, 0.30]	[0.04, 0.36]
N	988	988	988	988
	Elasticity of Failures With Respect to ENPV			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(5)	(6)	(7)	(8)
Log(ENPV)	0.188***	0.226***	0.225**	0.146***
	[0.09, 0.23]	[0.07, 0.28]	[0.02, 0.31]	[0.07, 0.18]
N	988	988	988	988

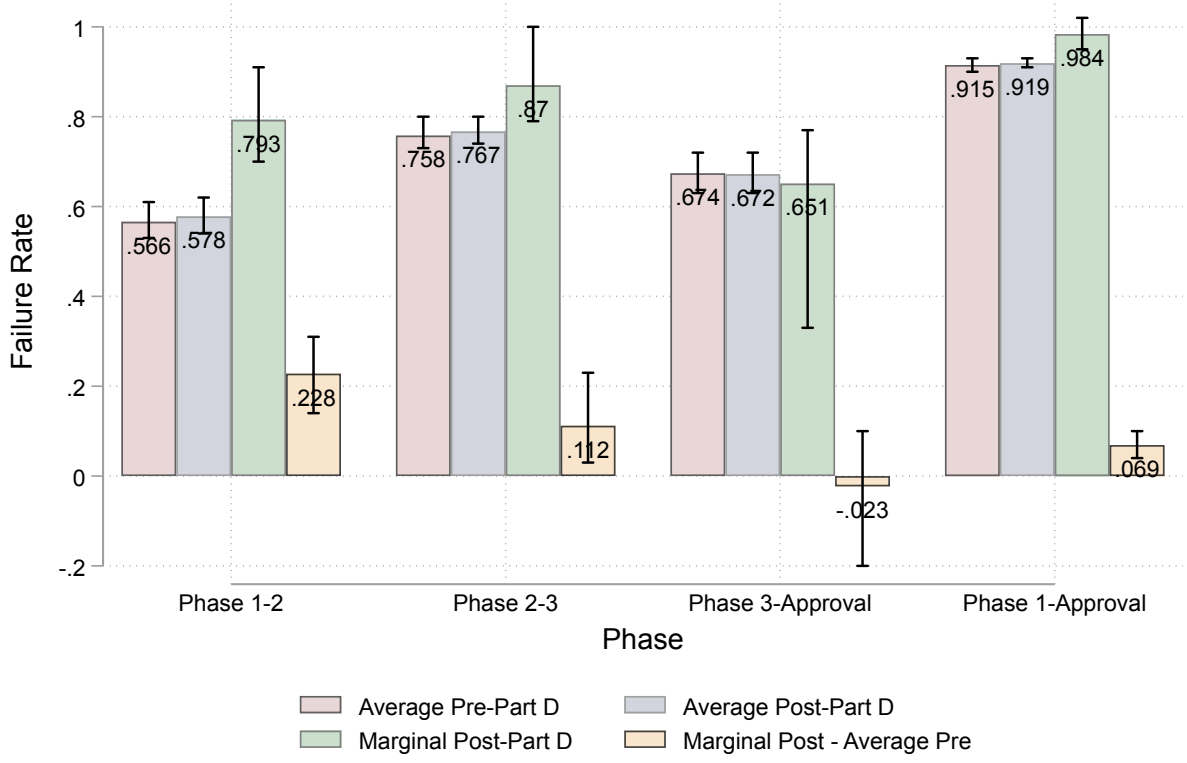
Notes: Data from PharmaProjects from 2001 to 2012. Bootstrap 95% confidence intervals, clustered by drug class, estimated with 5,000 iterations are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Failures Rates in Response to 50% Change in ENPV

	Failure Rate From Phase 1 to 2 (1)	Failure Rate From Phase 2 to 3 (2)	Failure Rate From Phase 3 to Approval (3)	Failure Rate From Phase 1 to Approval (4)
	Marginal Post-Part D Failure Rate			
Log(ENPV)	0.793***	0.870***	0.651**	0.984***
	[0.70, 0.91]	[0.79, 1.00]	[0.33, 0.77]	[0.95, 1.02]
N	988	988	988	988
	Average Post-Part D Failure Rate			
Log(ENPV)	0.578***	0.767***	0.672***	0.919***
	[0.54, 0.62]	[0.74, 0.80]	[0.63, 0.72]	[0.91, 0.93]
N	988	988	988	988

Notes: Data from PharmaProjects from 2001 to 2012. Bootstrap 95% confidence intervals, clustered by drug class, estimated with 5,000 iterations are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Figure 6. Average and Marginal Failure Rate Due to 50% Change in ENPV



Notes: Data from PharmaProjects from 2001 to 2012. Bootstrap 95% confidence intervals, clustered by drug class, estimated with 5,000 iterations are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Marginal and Average Out-of-pocket Costs to Develop One Drug

	Marginal Drug Post-Part D (1)	Average Drug Pre-Part D (2)	Average Drug Post-Part D (3)	Marginal Increase (1)-(2) (4)
	A. Number Needed for One Drug			
Phase I	106.48	29.21	31.01	77.27
Phase II	22.04	12.68	13.08	9.36
Phase III	2.87	3.07	3.07	-0.20
	B. Cost for One Drug ($2018 million)			
DiMasi et al. (2003)	3476.00	1467.76	1519.34	2008
Adams and Bratner (2006)	1269.84	866.46	880.79	403
DiMasi et al. (2016)	5084.84	2441.62	2511.43	2643
Sertkaya et al. (2024)	1474.73	747.48	767.19	727

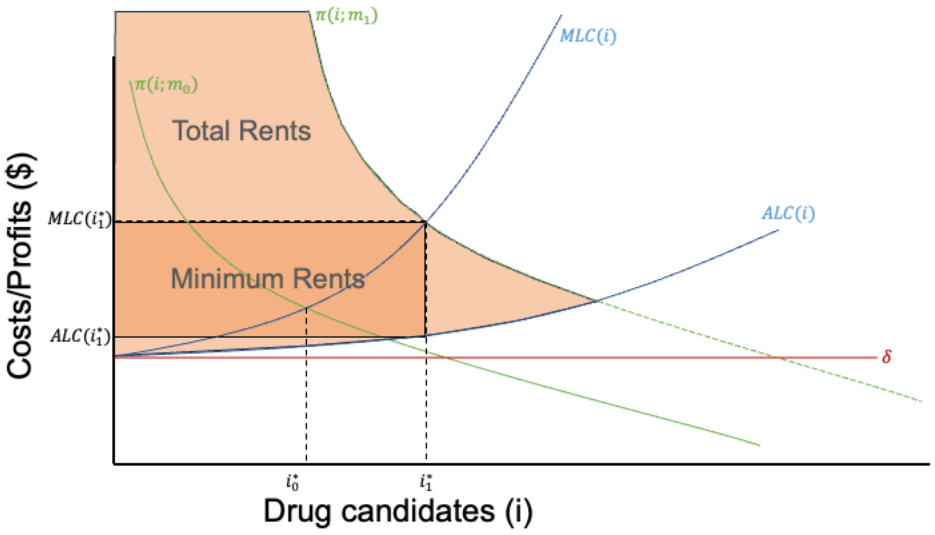
Notes: We use data from Table 4 to identify the number of phase I, II, and III trials needed to launch one drug (see **Error! Reference source not found.**). The costs per phase come from several different sources in the literature and are all expressed in $2018.

Table 6. Changes in the Quality Distribution of Drugs Induced by Part D

	Major, Significant, or Moderate Improvements	Minor Improvement	No Improvements
	A. Marginal Effect With and Without Part D		
	(1)	(2)	(3)
1(Part D) x MMS	0.116***	0.028	-0.144***
	[0.04, 0.20]	[-0.06, 0.09]	[-0.21, -0.07]
N	487	487	487
	B. Number of Drugs		
	(4)	(5)	(6)
1(Part D) x MMS	1.563***	0.514	-0.862**
	[0.36, 1.67]	[-0.39, 0.89]	[-1.32, -0.07]

Notes: Panel A estimates multinomial logit regressions examining changes in the probability of ASMR I-III, IV, and V. Panel B calculates the class-level average predicted probability of quality with and without Part D and multiplies it to the predicted number of new approvals with and without Part D, respectively. The resulting changes in ASMR 1-III, IV, and V drugs as a result of Part D is then calculated by summing the changes across classes.

Figure 7. The Implications of Upward-Sloping Innovation Supply for Pharmaceutical Rents.



Notes: The green curves depict $\boldsymbol{\pi(i;m)}$, the post-approval profits of a type $\boldsymbol{i}$ drug when market size is $\boldsymbol{m}$. The blue curves depict the marginal cost of launching a type $\boldsymbol{i}$ drug ($\boldsymbol{MLC(i)}$) and the average cost of launching $\boldsymbol{i}$ drugs ($\boldsymbol{ALC(i)}$). $\boldsymbol{\delta}$ is the cost of researching a drug candidate.

IX. Appendix A: Mathematical Derivations

Here we present the proof of Implication (1) from Section II.C.

Implication 1: If $\boldsymbol{D_i(i;m) < S_i(i)}$ then $\frac{di^*}{dm} > 0$.

Proof: Firms must decide which drug candidates to pursue. In equilibrium, firm decisions will maximize total producer surplus:

$$\max_{i^*} \int_0^{i^*} [(\sigma(i))\pi(i;m) - \delta]di \tag{8}$$

The risk level of the marginal drug satisfies the first-order condition:

$$\pi(i^*;m) = \frac{\delta}{\sigma(i^*)} \tag{9}$$

Defining SOC as the second-order condition, where $SOC \equiv \frac{\partial}{\partial i^*}[(\sigma(i^*))\pi(i^*;m)]$, comparative static arguments imply that:

$$\frac{di^*}{dm} = \frac{\sigma(i^*)\pi_m(i^*;m)}{-SOC} \tag{10}$$

Differentiation reveals:

$$SOC \equiv \sigma'(i^*)\pi(i^*;m) + \sigma(i)\pi_i(i^*;m)$$

Recall that $S(i) \equiv \frac{\delta}{\sigma(i)}$ and $D(i;m) \equiv \pi(i;m)$. Therefore, $S'(i) = -\frac{\delta}{\sigma(i)^2}\sigma'(i)$ and $D_i(i;m) = \pi_i(i;m)$. Therefore, we can rewrite SOC as:

$$SOC \equiv -\frac{S'(i^*)D(i^*;m)\sigma(i^*)}{S(i^*)} + \frac{\delta D_i(i^*;m)}{S(i^*)} \tag{11}$$

Multiplying both sides by $\frac{S(i^*)}{\delta} = \frac{1}{\sigma(i)}$, and recognizing that $D(i^*) = S(i^*)$ in equilibrium, this simplifies as:

$$\frac{SOC}{\sigma(i)} \equiv -S'(i^*) + D_i(i^*;m) < 0 \tag{12}$$

The inequality results from the assumption of the claim. This then implies the desired result:

$$\frac{di^*}{dm} = \frac{\sigma(i^*)\pi_m(i^*;m)}{-SOC} > 0 \tag{13}$$

X. Appendix B: Data Details

A. Identifying Drug Development Pipelines

For each drug ingredient, PharmaProjects records the furthest development statuses reached in any country, by any company, for any disease. Drugs in active development have dates for phases I, II, and III, and approval from relevant regulatory authorities. A drug becomes classified as "inactive" when it is suspended, discontinued, withdrawn, or missing records for a year. Thus, we treat trials which do not have subsequent development dates as having failed. If there is a missing date for a previous phase, we do not treat the missing observation as a failure. For example, a drug that has phase I and III dates, but is missing a phase II date, will have successfully progressed from phase I to II and phase II to III, and we count the drug as having had phase I, II, and III trials.

We show the duration of clinical trial developments from phase I to approval across different cohorts of drugs in Appendix Table B1.

Appendix Table B1. Duration of Clinical Development from Phase I to Approval

Phase I Cohort	25th Percentile	50th Percentile	75th Percentile
1996	5.77	8.42	12.02
1997	3.17	7.72	13.96
1998	6.02	7.32	13.81
1999	5.30	6.75	9.32
2000	6.66	9.38	10.68
2001	5.30	9.36	10.51
2002	6.45	9.57	13.75
2003	7.40	9.48	11.58
2004	5.35	7.70	10.99
2005	4.19	7.26	9.46
2006	4.76	8.19	10.04
2007	5.85	7.23	10.92
2008	5.73	7.13	9.04
2009	6.21	7.84	9.34
2010	5.22	6.64	8.39
2011	4.77	5.61	7.77
2012	3.86	6.16	7.37

Constructing the Therapeutic Crosswalk

To match the PharmaProjects therapeutic classes to the Multum therapeutic classes in MEPS, we relied on exact and fuzzy string matching. For strings that were not exact matches in PharmaProjects, we manually searched for the appropriate Multum class by examining class descriptions and comparing drugs in the Multum class with those in PharmaProjects. The table below documents the crosswalk we created. Over time, Multum therapeutic classes have grown in complexity. To the extent possible, we retain the 2001 to 2003 coding system since our analysis relies on the Medicare market share within each drug class between that time period.

Appendix Table B2. Matching of Multum Therapeutic Classes

Multum No.	Multum Therapeutic Class Description (MEPS)	PharmaProjects Therapeutic Class Descriptions
2	amebicides	amoebicide; protozoacide
3	anthelmintics	anthelmintic; parasiticide; schistosomicide
4	antifungals	antifungal
5	antimalarial agents	antimalarial
6	antituberculosis agents	antimycobacterial
9	cephalosporins	cephalosporin, injectable; cephalosporin, oral
11	macrolides derivatives	macrolide antibiotic
12	miscellaneous antibiotics	antibiotic, other; peptide antibiotic; trimethoprim and analogues
13	penicillins	penicillin, injectable; penicillin, oral
14	quinolones	quinolone antibacterial
16	tetracyclines	tetracycline antibiotic
18	aminoglycosides	aminoglycoside antibiotic
19	antihyperlipidemic agents	hypolipaemic/antiatherosclerotic, miscellaneous antihyperlipidemic agents
21	alkylating agents	anticancer, alkylating
22	antineoantibiotic	anticancer, antibiotic
23	antimetabolites	alimentary/metabolic, other; anticancer, antimetabolite
24	hormones/antineoplastics	anticancer, hormonal
25	miscellaneous antineoplastics	antisickling
31	bacterial vaccines	therapeutic vaccine, anti-infective
32	colony stimulating factors	cytokine
36	recombinant human erythropoietins	antianaemic; recombinant growth factor; recombinant hormone; recombinant interleukin; recombinant vaccine; recombinant, other
38	viral vaccines	prophylactic vaccine, anti-infective; vaccine adjuvant
43	antiadrenergic agents, peripherally acting	antihypertensive, adrenergic
43	antiadrenergic agents, peripherally acting	prostate disorders

45	antianginal agents	antianginal
46	antiarrhythmic agents	antiarrhythmic
51	miscellaneous cardiovascular agents	antithrombotic, fibrinolytic; antithrombotic, other; cardiovascular
52	peripheral vasodilators	vasodilator, peripheral
53	vasodilators	vasodilator, coronary; vasodilator, renal
55	antihypertensive combinations	antihypertensive, diuretic; antihypertensive, other; antihypertensive, renin system; antihypotensive
58	analgesics	analgesic, nsaid; analgesic, opioid; analgesic, other; anti-inflammatory; anti-inflammatory, topical; antimigraine
64	anticonvulsants	antiepileptic
65	antiemetic/antivertigo agents	antiemetic
66	antiparkinson agents	antiparkinsonian; neurological
67	anxiolytics, sedatives, & hypnotics	anxiolytic; hypnotic/sedative
71	CNS stimulants	cardio stimulant; psychostimulant
73	muscle relaxants	multiple sclerosis treatment; muscle relaxant; musculoskeletal
82	anticoagulant	antithrombotic, anticoagulant
88	antacids	antacid/antiflatulent
89	anticholinergics/antispasmodics	antispasmodic
90	antidiarrheals	antidiarrheal
91	digestive enzymes	digestive
92	gallstone solubilizing agents	gallstone therapy
93	gi stimulants	gastroprokinetic
95	laxatives	laxative
99	antidiabetic agents	antidiabetic; symptomatic antidiabetic
100	miscellaneous hormones	glucagon; hormone; releasing hormone; somatostatin
101	sex hormones	anabolic; female contraceptive; female sexual dysfunction; fertility enhancer; labour inhibitor; male contraceptive; menopausal disorders; menstruation disorders
103	thyroid drugs	antithyroid; thyroid hormone
104	immunosupressive agents	immunological, unspecified; immunomodulator, anti-infective; immunosuppressant
106	antidotes	antidote; hepatoprotective
110	miscellaneous uncategorized agents	immunostimulant, other
112	radiocontrast agents	imaging agent; radio/chemoprotective; radio/chemosensitizer
113	genitourinary tract agents	male sexual dysfunction; reproductive/gonadal, general; urological
119	vitamins	nutritional supplement
121	intravenous nutritional products	metabolic and enzyme disorders
123	antihistamines	antiallergic, non-asthma
123	antihistamines	antipruritic/inflamm, allergic
124	antitussives	antitussive
125	bronchodilators	copd treatment

129	miscellaneous respiratory agents	respiratory
131	antiasthmatic combinations	antiasthma
135	antiseptic and germicides	septic shock treatment
136	dermatological agents	anaesthetic, inhalation; anaesthetic, injectable; anaesthetic, local; antiacne; dermatological
147	ophthalmic preparations	ophthalmological, antiglaucoma; ophthalmological, other; plasma substitute
148	otic preparations	otological
192	antirheumatics	antiarthritic, immunological
194	antigout agents	antigout
249	antidepressants	antidepressant; antineurotic
251	antipsychotics	antipsychotic
253	anorexiants	anorectic/antiobesity
256	interferons	antiviral, interferon
257	monoclonal antibodies	antibody, non-monoclonal; immunoconjugate, other; monoclonal antibody, chimaeric; monoclonal antibody, human; monoclonal antibody, other
270	antipsoriatics	antipsoriasis
272	proton pump inhibitor	antiulcer
284	viscosupplementation agents	antiarthritic, other
293	insulin syringes	insulin
295	growth hormones	growth hormone
313	cholinesterase inhibitors	cognition enhancer
320	smoking cessation agents	dependence treatment
323	antineoplastic monoclonal antibodies	anticancer, immunological
324	antineoplastic interferons	anticancer, interferon
343	tyrosine kinase inhibitors	anticancer, other; anticancer, vaccine

B. Matching PharmaProjects and MEPS

Pharmaceutical revenues in MEPS come from household-reported prescription drug use. Validation studies have found that total expenditures on drugs were close to the Part D claims data, both on average and across the distribution of expenditures (Hell, Zuvekas, and Zodet 2011). As a result, we are unable to capture drug-class revenues for non-Part D drugs. PharmaProjects follows all drug development, including physician-administered infusions reimbursed under Part B. In Appendix Table B3, we show all unmatched therapeutic classes in PharmaProjects where there were more than 20 drugs in development between

1996 and 2012. The table demonstrates that unmatched therapeutic classes were indeed physician-administered drugs.

Appendix Table B3. Therapeutic Classes in PharmaProjecs Without MEPS Revenues

TC1S1	No. of Drugs	Therapeutic Class Description	PharmaProject Drug Examples
323	750	antineoplastic monoclonal antibodies	"mabs"
343	671	tyrosine kinase inhibitors	"nibs"
112	120	radiocontrast agents	florbetaben, pegfilgrastim, lipegfilgrastim
131	110	antiasthmatic combinations	none approved
36	62	recombinant human erythroproteins	epoetin alfa, galsuflase
19	56	antihyperilic agents	ezetimibe and simavstatin, lomitapide
106	51	sulfonamides	alipogene tiparvovec, netakimab, tolvaptan
38	46	viral agents	influenza vaccine
284	33	vicosupplementation agents	none approved
129	24	miscellaneous respiratory agents	vismodegib
21	21	alkylating agents	bendamustine

XI. Appendix C: Additional Tables and Figures

Appendix Table C1 OLS Estimates of ENPV Increases on Clinical Trials, Failures, and Marginal Failure Rates

	Trials			
	Phase I Trials	Phase II Trials	Phase III Trials	Approvals
	(1)	(2)	(3)	(4)
Log(ENPV)	1.239***	0.454***	0.203***	0.090**
	[0.37, 2.30]	[0.10, 0.84]	[0.05, 0.43]	[0.02, 0.17]
N	1,035	1,035	1,035	1,035
	Failures			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(5)	(6)	(7)	(8)
Log(ENPV)	0.812***	0.367**	0.121*	1.149***
	[0.25, 1.54]	[0.08, 0.63]	[-0.02, 0.29]	[0.31, 2.15]
N	1,035	1,035	1,035	1,035
	Marginal Failure Rates			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(9)	(10)	(11)	(12)
Log(ENPV)	0.655***	0.809***	0.598*	0.928***
	[0.54, 0.81]	[0.55, 1.16]	[-0.18, 0.92]	[0.80, 0.98]
N	1,035	1,035	1,035	1,035

Notes: Data from PharmaProjects from 2001 to 2012. Second-stage results use estimate an ordinary least squares regression with drug class fixed effects. Bootstrap 95% confidence intervals, clustered by drug class, estimated with 5,000 iterations are shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table C2.Impact of Varying Trial Follow-Up Duration

	Phase I	Phase II	Phase III	Approvals
	(1)	(2)	(3)	(4)
	Elasticity of Trials Within 10 Years of Phase I Start			
Log(ENPV)	0.170***	0.196***	0.171**	0.175*
	[0.06, 0.23]	[0.03, 0.29]	[0.00, 0.30]	[-0.00, 0.33]
N	1,035	1,035	1,035	1,035
	Elasticity of Trials Within 14 Years of Phase I Start			
Log(ENPV)	0.170***	0.196***	0.178**	0.184**
	[0.06, 0.23]	[0.04, 0.29]	[0.01, 0.30]	[0.00, 0.33]
N	1,035	1,035	1,035	1,035
	Elasticity of Failures Within 10 Years of Phase I Start			
Log(ENPV)	0.209***	0.216***	0.157*	0.182***
	[0.08, 0.28]	[0.05, 0.32]	[-0.03, 0.30]	[0.07, 0.24]
N	1,035	1,035	1,035	1,035
	D. Elasticity of Failures Within 14 Years of Phase I Start			
Log(ENPV)	0.208***	0.215***	0.163*	0.180***
	[0.08, 0.28]	[0.05, 0.32]	[-0.01, 0.30]	[0.07, 0.24]
N	1,035	1,035	1,035	1,035

Notes: Generalized linear model with a Poisson distribution and log link estimated. Bootstrap 95% confidence intervals clustered by drug class are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table C3.Effect among manufacturers without drugs on the market before Part D

	Trials			
	Phase I Trials	Phase II Trials	Phase III Trials	Approvals
	(1)	(2)	(3)	(4)
Log(ENPV)	0.111***	0.148***	0.246***	0.327***
	[0.05, 0.15]	[0.07, 0.20]	[0.13, 0.33]	[0.18, 0.40]
N	1,035	1,035	1,035	1,035
	Failures			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(5)	(6)	(7)	(8)
Log(ENPV)	0.159***	0.175***	0.284***	0.117***
	[0.08, 0.20]	[0.09, 0.22]	[0.15, 0.35]	[0.06, 0.15]
N	1,035	1,035	1,035	1,035
	Marginal Failure Rates			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(9)	(10)	(11)	(12)
Log(ENPV)	0.826***	0.883***	0.762***	0.972***
	[0.67, 1.11]	[0.75, 1.04]	[0.59, 0.91]	[0.90, 1.04]
N	1,035	1,035	1,035	1,035
	Average Failure Rates			
	Failures Phase I to II	Failures Phase II to III	Failures Phase III to Approval	Failures Phase I to Approval
	(13)	(14)	(15)	(16)
Log(ENPV)	0.576***	0.751***	0.658***	0.918***
	[0.51, 0.69]	[0.71, 0.85]	[0.58, 0.79]	[0.89, 0.95]
N	1,035	1,035	1,035	1,035

Notes: Generalized linear model with a Poisson distribution and log link estimated. Bootstrap 95% confidence intervals clustered by drug class are in brackets. *** p<0.01, ** p<0.05, * p<0.1.