Can Financial Engineering Cure Cancer?:
A New Approach for Funding Large-Scale Biomedical Innovation

Jose-Maria Fernandez\textsuperscript{1}, Roger M. Stein\textsuperscript{1,2}, Andrew W. Lo\textsuperscript{1,3,4}

\textsuperscript{1}MIT Sloan School of Management and Laboratory for Financial Engineering, 100 Main Street, E62–618, Cambridge, MA 02142, USA
\textsuperscript{2}Moody’s Corporation, 7 World Trade Center, New York, NY 10007, USA
\textsuperscript{3}MIT CSAIL and EECS, 32 Vassar Street, 32–G608, Cambridge, MA 02139, USA
\textsuperscript{4}AlphaSimplex Group, LLC, One Cambridge Center, Cambridge, MA 02142, USA

Biomedical research is an expensive, lengthy, and risky process that challenges traditional funding vehicles which are limited in size, scope, and risk appetite. Here we propose a new funding model that uses financial engineering techniques to raise large amounts of capital through debt and equity markets to be invested by a single fund in a well-diversified portfolio of biomedical research programs. Having a single entity invest in many diverse programs is a new business model that increases the likelihood of success to the point where the risk/reward profile becomes sufficiently attractive to a large population of investors. We construct an analytical framework to highlight the mechanism by which “megafunds” can greatly reduce the risk of innovation through diversification, allowing such entities to raise much larger pools of capital. In a hypothetical simulation using historical data for cancer drug-development programs, we find that megafunds of $5 to $15 billion are capable of yielding average investment returns in the range of 8.9\% to 11.4\% for equityholders, and 5\% to 8\% for bondholders. Though these returns may not satisfy venture-capital hurdle rates, they are attractive to pension funds, insurance companies, and other large institutional investors. Through closer collaboration between scientists, clinicians, and financial engineers, megafunds may become a practical method for funding large-scale biomedical innovations.
Background

There is a growing consensus that the “bench-to-bedside” process of translating biomedical research into effective therapeutics is broken. A confluence of factors is responsible for such pessimism but among the most widespread is the sense that the current business model for life sciences research and development is flawed. The productivity of big pharmaceutical companies—as measured by the number of new molecular entities and biologics license applications per dollar of R&D investment—has declined in recent years, and their stock-price performance over the last decade—an annualized return of $1.2\%$ for the NYSE/ARCA Pharma Index during the period from January 2, 2002 to January 4, 2012—has been equally disappointing. Despite the fact that the aggregate R&D budget of the pharmaceutical industry has almost doubled from $68$ billion in 2002 to $127$ billion in 2010, there has been little appreciable impact on the number of new drugs approved. Life sciences venture-capital investments have not fared much better, with an average internal rate of return (IRR) of $-1\%$ over the 10-year period from 2001 through 2010 according to VentureXpert data. Against a backdrop of declining real prescription-drug spending, rising costs, shrinking R&D budgets, expiring blockbuster patents, post-Vioxx fallout, lower levels of funding and risk tolerance among venture capitalists, and unprecedented stock-market volatility and uncertainty, it is not surprising that the future of this industry appears so bleak. However, this outlook seems to contradict the many promising breakthroughs that have occurred in biomedicine in recent years, including successful stem-cell therapies such as bone marrow transplantations, powerful new computational tools for medical imaging and radiosurgery, diagnostic applications of nanotechnology, the identification of biomarkers for certain diseases such as prostate cancer and
heart disease, and, of course, the sequencing of the human genome. Moreover, there are many life-threatening diseases for which the number of afflicted individuals continues to increase—if for no other reason than due to population growth—implying a growing demand for therapeutics from a grateful and price-insensitive clientele. Why, then, does the industry appear to be so challenged?

Here we propose one explanation for this contradiction and a possible solution. Biomedical research is complex, expensive, uncertain, lengthy, and fraught with conflicting non-pecuniary motivations and public-policy implications. While other industries may share some of these characteristics, it is difficult to find another so heavily burdened by all of them. These traits suggest that public and private equity may not be the most effective funding sources for such endeavors because the needs and expectations of shareholders and limited partners are fundamentally inconsistent with the realities of biomedical innovation. The traditional real-time pricing and dispersed ownership of public equities implies constant scrutiny of corporate performance from many different types of shareholders, pushing senior management toward projects with clearer and more immediate payoffs, and away from more speculative but potentially transformative R&D. Private equity may afford more latitude for risk-taking and deferred exits, but the scale of capital commitment is considerably smaller and funding cycles are driven not as much by scientific breakthroughs as they are by stock-market and business cycles that determine when limited partners can cash out. Recent theoretical and empirical evidence suggest that even the mere concern about the availability of future rounds of financing—due solely to the possibility of unfavorable economic conditions—is often reason enough for venture capitalists (VCs) to shun proven and economically viable technologies. Industry participants cite the existence of a “valley of death”—a funding
gap between basic biomedical research and clinical development. For example, in 2010 only $6 to $7 billion was spent on translational efforts while $48 billion was spent on basic research and $125 billion was spent on clinical development that same year.\textsuperscript{7}

We propose an alternative for funding biomedical innovation that addresses these issues through the use of “financial engineering”\textsuperscript{8,9} mathematical and statistical models for structuring and pricing various financial securities to achieve specific objectives. Our approach involves two components: (1) creating large diversified portfolios—on the order of $5 to $15 billion—of projects; and (2) structuring the financing for these portfolios as combinations of equity and securitized debt so as to access much larger sources of investment capital.

The ability to create such “megafunds” has the potential to radically alter the existing business model of the biopharma industry. A megafund can extend the investment horizon and reduce the financing risk for the programs in its portfolio, allowing research to progress along the most scientifically productive paths, less encumbered by short-term business pressures. In addition to the benefits of diversification, running many concurrent research programs through a single organization may also yield greater operating efficiency. These gains are achieved not only through the usual economies of scale, but also through the ability to discontinue less promising projects sooner and to redeploy researchers onto more productive projects.

For the plethora of small pharmaceutical companies pursuing one or two projects, these savings are especially significant. Despite the limited number of compounds they own, these small companies must pay the same fixed costs as larger entities to comply with Food and Drug Admin-
istration (FDA) regulations and, if they are public, SEC rules and the Sarbanes-Oxley Act. Much of these nontrivial duplicative costs would be eliminated if a large number of compounds were developed simultaneously by a single megafund. Also, it is considerably harder to cull compounds efficiently in a small company because the livelihoods of the employees and management depend on the continued development of the company’s few compounds. In these cases, development tends to continue until the money runs out. With a megafund, this conflict is greatly reduced—capital can be more efficiently allocated to projects that are likely to succeed while failing projects and compounds can be abandoned rapidly.

Of course, the organizational complexities of managing a portfolio of highly heterogeneous biomedical projects also increase with scale, which can reduce some of the benefits of diversification. These dis-economies of scale suggest that the optimal size of a megafund may differ from one application to the next, and will be determined by balancing organizational complexity against scientific, operational, and financial synergies. New approaches to management, corporate governance, and scientific collaboration may also be necessary before significantly larger amounts of capital can be profitably deployed in this industry. These important issues lie beyond the scope of this paper, but our analysis suggests that if such issues can be resolved, new financing techniques can greatly expand the current scale of biomedical innovation.
Portfolio Theory

To illustrate our approach, we first present a highly simplified example here, and then describe a more detailed case study below using historical oncology drug-development data.

Consider a hypothetical biomedical project requiring $200 million in out-of-pocket costs over a 10-year period during which no revenues are generated. Suppose that the program’s probability of successfully developing a blockbuster drug is 5%. Few investors outside the biopharma sector would be tempted by such an opportunity, even though the expected rate of return on this investment may be quite attractive. In fact, if a blockbuster drug is assumed to generate net income of $2 billion per year over a 10-year period of exclusivity from years 11 to 20, the present value in year 10 is $12.3 billion (using a 10% cost of capital$^{10}$), implying an expected compound annual rate of return of $11.9\% = (0.05 \times 12.3/0.2)^{1/10} - 1$ over the 10-year investment period (see Figure 1 for the timeline of these cashflows). However, investors do not earn 11.9% with certainty, but face two possible outcomes instead: a 95% probability of earning $-100\%$ and a 5% probability of earning $51.0\% = (12.3/0.2)^{1/10} - 1$. These projects may simply be too risky for most investors given the near certainty of getting wiped out and the long wait before any revenues are generated.

However, consider investing in 150 such programs simultaneously through a single investment vehicle with $150 \times 200$ million $= 30$ billion of investable capital, which we shall refer to as a “megafund”. For simplicity, assume that the success or failure of each program is a statistically independent draw. Then the probability of at least one success among 150 independent programs is $99.95\% = 1 - 0.95^{150}$, which is quite a different proposition. Although the expected profit of each
of the 150 programs remains the same at $12.3 billion, the likelihood of at least one hit is dramatically increased, reducing the risk of the entire portfolio. One simple measure of this risk reduction is the standard deviation of the annualized return, which is 423% for an individual draw, but only $34.6\% = \frac{423\%}{\sqrt{150}}$ for the annualized portfolio return. The more risky and less correlated the underlying assets are, the greater the benefits to pooling them, not unlike an insurance pool that provides protection for each of its participants by spreading any given individual’s losses over the entire membership. Such pools become more effective as the number of participants increases, and the same is true for megafunds of drug-development projects.

This risk reduction is not costless, but comes at the expense of a much greater capital commitment. Also, unless the individual assets in a portfolio are mutually uncorrelated (which is exceedingly improbable), modern portfolio theory\textsuperscript{11} shows that there is a limit to the amount of risk that can be eliminated through diversification. This limit and the optimal size of the megafund depend on several factors including the pairwise correlations between the assets’ returns, the natural scale of the investment in each asset, and the risk appetite, expected-return requirements, and investable wealth of the population of potential investors. While some investors may prefer the high-risk/high-return profile of a one-shot drug-development program, a much larger pool of investors seem to prefer the lower-risk/lower-return profile of a portfolio of programs, as demonstrated by the relative sizes of the venture capital industry ($176$ billion) and the mutual fund industry ($11.8$ trillion).\textsuperscript{12}

This example highlights the most important difference between a megafund and existing
biopharma financing alternatives: the risk reduction from diversification allows the megafund to issue large amounts of public debt and equity, greatly broadening the pool of potential investors willing to fund such projects. To see why, suppose each of the 150 projects was undertaken by a separate company, yielding 150 companies with development costs of $200 million apiece. The all-or-nothing nature of each company’s payoff implies that even if a company issued only a small amount of debt, the default probability of such bonds would be 95%. With default nearly guaranteed, debt financing is virtually impossible for these single-project entities, and the riskiness of a single project implies that an initial public offering of public equity is also unlikely.

However, a single entity with 150 such programs could issue $24.6 billion of zero-coupon bonds—bonds that pay only one lump-sum payment at maturity—maturing in year 10 with a default probability of only 0.4% (the probability of less than 2 successes, since two hits yield a present value of $24.6 billion in year 10, just enough to pay off the bondholders). This default probability is comparable to the historical realized 10-year default rates of the highest-rated category of debt (Aaa) from 1920 to 2010 according to the bond-rating agency Moody’s.13 As of February 2012, Moody’s reported the average yield of seasoned Aaa corporate bonds with approximately 30 years to maturity to be 3.85%,14 which is a reasonable proxy for the yield of a 10-year bond with identical credit quality. At a yield of 3.85%, a zero-coupon bond that promises to pay $24.6 billion in year 10 would generate proceeds of $16.8 billion when issued in year 0. If the remaining $13.2 billion were financed by equity, the expected rate of return and standard deviation would be 21.5% and 78.9%, respectively. These values are higher than those of the all-equity-financed case (11.9% and 34.6%) because of leverage, but are still within the range of risk/reward
profiles of publicly traded equities. A megafund’s ability to issue both debt and equity with attractive terms to a larger pool of potential investors provides greater scale and diversification benefits, yielding greater risk reduction and bigger overall impact on biomedical innovation.

The lower-risk/lower-return profile of a megafund may have little appeal to VCs, but is of much greater interest to pension funds, insurance companies, money market funds, banks, and other large financial institutions, who control a vastly larger pool of investment capital than VCs. For example, at the end of 2010 the California Public Employees Retirement System held $226 billion of investable assets, the Norwegian government pension fund held $537 billion, and non-government U.S. institutional money market funds held $1.1 trillion. Moreover, as of the end of 2010, the total size of the U.S. bond market was $35.2 trillion.

Of course, we have grossly oversimplified the economics of the biopharma industry in this example to highlight the mechanism by which investment performance can be enhanced through diversification. Raising large amounts of capital may seem unrealistic, especially given recent corporate consolidations, budget cutbacks, and capital scarcity. We address this issue below in the second component of our framework: securitization.

Securitization

Given the scale of financing needed for creating a truly diversified portfolio of drug-development investments and the time lag between capital commitment and return, private-partnership structures such as venture capital may not be ideal sources of capital for this industry. Instead, we propose
tapping public capital markets directly via “securitization”, a financial-engineering technique in which large amounts of investment capital are obtained from a diverse investor population by issuing debt and equity securities which are claims on a portfolio of assets—in our case, biomedical research. This technique figured prominently in the build-up to the recent financial crisis in which trillions of dollars of investment capital were raised and channeled into the U.S. residential real estate market. While there is still considerable controversy as to the ultimate causes of the crisis and the necessary reforms, there is little doubt that securitization is an effective means of raising capital.

A common form of securitization involves “cashflow” transactions in which a portfolio of assets—typically mortgages, auto loans, student loans, or credit-card receivables—is acquired using money raised by issuing equity and bonds of different seniorities. These assets and the cash-flows they generate are pledged as collateral for the debt (see Figure 2). In our proposed application, the assets include the initial capital raised from investors, all the subsequent biomedical R&D acquired or licensed from external entities or created internally, and all the profits generated by these activities or through sales of these assets in later periods. We shall refer to debt that is collateralized by such assets as “research-backed obligations” (RBOs).

To ensure that the portfolio of assets is used only to satisfy the payments of the newly issued RBOs, a stand-alone legal entity called a “special purpose vehicle” (SPV) is formed for the express purpose of purchasing the collateral and issuing and servicing the securities. Equityholders own equity in the SPV and thus have a claim on the residual assets and cashflows that remain after all
debt obligations have been satisfied. The SPV is managed by a separate management company, though for the purposes of this exposition, we may consider the SPV as both the megafund and the corporate entity that structures the biomedical R&D and external acquisitions and licensing deals.

To provide different levels of risk and expected return for the broadest possible set of potential investors, the RBOs are divided into distinct classes or “tranches” with different repayment priorities. The “senior” tranche has highest priority, meaning that in each payment period its obligations must be satisfied first before those of any other tranche, and each of the more “junior” tranches are repaid in order of their priority. In the event that the assets do not generate sufficient cashflow to make all promised payments to bondholders in any given period, the senior-most tranche will be paid first, followed by the next most senior tranche and so forth, until the available cash is exhausted. Therefore, the senior tranche is the least likely to experience losses; hence, it will carry the lowest but safest yield, which appeals to the most risk-sensitive investors such as money-market funds, banks, and smaller pension funds. More junior tranches have higher loss probabilities and must offer correspondingly higher yields to compensate investors for this increased risk, which attracts more risk-tolerant investors such as large pension funds, endowments, and high-net-worth private investors. The most junior tranche is often structured as equity—and sometimes called the “equity tranche”—with no promised payments whatsoever but with unlimited upside potential once bondholders are repaid in full. Equityholders stand to reap the biggest gains if the megafund’s underlying assets do well, but they are the first to suffer losses if those assets are not profitable. As a result, this is the riskiest tranche and most likely to be purchased by the most risk-tolerant portion of the investor population, i.e., hedge funds, funds of funds, VCs, and
deep-pocketed institutional investors including big pharmaceutical companies. The size and order of the tranches is known as the SPV’s “capital structure” and the motivation for multiple tranches should now be clear: regardless of how risk-averse or risk-seeking an investor is, there is likely to be a particular tranche of this SPV’s debt issue that will satisfy the investor’s risk preferences.

In addition to the different levels of priority, the RBOs can be customized in several important ways. The maturity of the RBOs can be extended so as to eliminate the shorter-term pressures of generating earnings and preparing for an initial public offering, which can often lead to the distressed sale of promising but early-stage assets. Typical securitizations employ debt maturities of 15 years or less; for example, in August 2007 Drug Royalty Corp. issued $356 million of 8- and 15-year bonds backed by major royalty rights to the FDA-approved biopharmaceutical products of Enbrel, Remicade, Preotact, and FluMist. Even longer-term straight-debt issues have been successfully launched more recently, including $750 million in 100-year bonds issued by the Massachusetts Institute of Technology in May 2011 at the historically low rate of 5.623%. Such long horizons contrast sharply with the considerably shorter horizons of VCs and other traditional funding vehicles. This is especially relevant for large-scale biopharma investments that require lengthy and often unpredictable R&D and clinical-trial processes, and in which random interruptions due to lack of funding almost always destroy significant economic value, even for genuinely effective therapeutics.

Additional features known as “credit enhancements” and “triggers” are often used to provide further protection for the RBOs’ most senior tranches. These features include default insurance
through credit-default swaps (CDS), over-collateralization, the use of interest- and debt-coverage ratio thresholds that trigger accelerated payments when breached, and government programs such as those offered by the Federal National Mortgage Association (Fannie Mae) and the Federal Home Loan Mortgage Corporation (Freddie Mac) for mortgage-related securitizations (see SI for further discussion). While the role of government agencies in the recent financial crisis is still being debated, there is broad consensus that the programs they instituted were quite effective in promoting homeownership. Similar programs could be implemented by government agencies and government sponsored enterprises for supporting biomedical megafunds.

The SPV’s capital structure, priority of payments, and various coverage tests and credit enhancements are collectively known as the “cashflow waterfall”—a reference to the manner in which cashflows from the SPV’s assets spill over from senior to junior tranches—which fully determines the financial structure of each of its corresponding securities and how investors will be compensated in all circumstances. Once the SPV’s cashflow waterfall is specified, the economic value of the securities issued by the SPV can be directly related to the performance of the SPV’s assets. If the statistical properties of those assets’ cashflows can be quantified, the risk/reward profile of the SPV can be estimated, its securities can potentially be rated by bond-rating agencies, and they can be evaluated and purchased by a broad universe of investors. Therefore, one of the key factors in determining whether a pool of assets can be securitized is the degree to which the stochastic properties of the underlying assets’ returns over time can be measured and managed. In the multi-trillion-dollar mortgage-backed securities market, the answer was (and still is) yes, as is the case for corporate debt and several other asset classes.17 We believe the same is true for biomedical
research. By creating a large portfolio of well-diversified biopharma investments—on the order of $15 billion—and by spreading the risks and rewards of such a portfolio across a much larger and more diverse group of investors through securitization, it may be possible to facilitate large-scale and long-term biomedical innovation in a sustainable and, ultimately, profitable manner.

Curing Cancer: An Illustrative Example

To illustrate the practicality of megafund financing, we present a detailed simulation a hypothetical funding vehicle for cancer drug-development programs. Our focus on cancer is motivated by four considerations. First, the high cost and long development cycle of cancer therapeutics make this disease group a natural application of megafund financing. Second, the lifetime probability of developing cancer in the U.S. is 1 in 2 for men and 1 in 3 for women,\textsuperscript{18} and the number of deaths caused by cancer worldwide will grow to over 12 million per year by 2030,\textsuperscript{19} creating an urgency and visibility that will greatly facilitate fundraising for a cancer megafund. Third, because cancer is a complex conjoint of over 200 diseases, the multiple approaches to anticancer therapies yield greater opportunities for portfolio diversification, offsetting to some degree the megafund’s singular focus on cancer. And finally, several databases of cancer drug-development programs are readily available, allowing us to construct more realistic simulations of the possible risks and returns from a cancer megafund. These simulations are critical for communicating the megafund’s risks and rewards to potential investors, a prerequisite for any successful fundraising effort.

However, these motivations must be tempered by the caveat that a megafund devoted solely
to cancer is likely to understate the benefits of diversification and megafund financing for at least two reasons: the unavoidable correlation among cancer drug-discovery programs due to common biochemical pathways and pathologies, and the fact that cancer drug-approval rates are the lowest among all therapeutic areas since 2004 (6.7% in oncology vs. 12.1% in all other areas combined as of 2011\textsuperscript{20}). A more effective approach would be to target many diseases in addition to cancer so as to increase diversification. Moreover, our simulation focuses exclusively on the development of anti-cancer compounds, which ignores several other important cancer therapies such as diagnostic tools, radiosurgery, and gene therapy for which we have much less historical data to draw on. As with any simulation, each of our parameters can be challenged as being too conservative or too optimistic, and our hypothetical business structure may be viewed as too simplistic. We acknowledge these concerns at the outset and encourage readers to experiment with our simulation software using their own calibrations (our complete source code is available in both R and Matlab under an open-source license that enables all researchers to use, modify, and distribute it).

For concreteness, the financing mechanism we consider in this illustration relies on the securitization of experimental drug compounds only, and the objective of the SPV would be to finance the development of each of its compounds while satisfying the megafund’s obligations to its bondholders and providing attractive returns to its equity investors. The business structure of the SPV is illustrated in Figure 3, and the types of payments made by the SPV during its life include:

1. **Start-Up Expenses and Purchases.** The SPV will deploy its initial capital by acquiring economic rights to anticancer compounds in exchange for upfront and milestone payments
as well as funding R&D and clinical trials.

2. **Initial Post-Launch Expenses and Principal and Interest.** Because it may take a number of years before its start-up activities begin generating revenues, the SPV will set aside an initial cash reserve to fund clinical trials for its portfolio of compounds during the life of the transaction. These reserves will also ensure that timely payments of interest can be made on the RBOs in the early stages of the megafund.

3. **Ongoing R&D and Financing Expenses.** The SPV will pay for ongoing R&D expenses during the life of the megafund. As part of this process, the SPV may decide to sell some of its assets and engage in other corporate transactions to realize gains, to meet funding needs, or for strategic reasons.

4. **Management Costs.** During each year, the SPV will pay salaries to its staff, fees to external service providers, and other operating costs that are part of the “management fee”, which is typically assessed as a fixed percentage of the SPV’s total assets under management.

5. **Sale of Portfolio.** Upon the maturity of the longest-dated RBO bonds, the SPV portfolio will be liquidated and the proceeds paid out to the equityholders.

Our analysis involves simulating the revenues and costs in each period during the life of the SPV as compounds advance through the R&D and approval process. We use historical industry values that are summarized in Table 1 and derived from various research studies and data from Bloomberg.\(^3\)\(^{-23}\) To calibrate the simulation of the clinical-trials process, we merged two datasets: the DEVELOPMENT optimizer™ database provided by Deloitte Recap LLC and a data
set provided by the Center for the Study of Drug Development (CSDD) at Tufts University School of Medicine. The merged data contained over 2,000 compounds which, after removing duplicates and compounds for which there was not enough information, yielded a final set of 733 new molecular entities developed primarily for anticancer indications that entered clinical trials between January 1990 and January 2011. These compounds were developed by either biotechnology or pharmaceutical companies and were either therapeutic compounds or vaccines (summary statistics for the data are provided in SI). Using this data and the results of Paul et al. (2010), we define seven distinct phases of drug development—the initial preclinical phase (PreC), the three stages of the clinical-trials process (Phases I, II, and III), new drug application (NDA), approval (APP), and withdrawal (WD)—and estimate the following transition probability matrix \( P \) using standard methods (see SI):

\[
P = \begin{pmatrix}
    \text{PreC}_{t+1} & \text{Phase I}_{t+1} & \text{Phase II}_{t+1} & \text{Phase III}_{t+1} & \text{NDA}_{t+1} & \text{APP}_{t+1} & \text{WD}_{t+1} \\
    0.50 & 0.35 & 0.00 & 0.00 & 0.00 & 0.00 & 0.15 \\
    0.00 & 0.80 & 0.14 & 0.01 & 0.00 & 0.00 & 0.05 \\
    0.00 & 0.00 & 0.86 & 0.06 & 0.00 & 0.00 & 0.08 \\
    0.00 & 0.00 & 0.00 & 0.85 & 0.07 & 0.02 & 0.06 \\
    0.00 & 0.00 & 0.00 & 0.00 & 0.56 & 0.40 & 0.04 \\
    0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 \\
    0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00
\end{pmatrix}
\]

Using this transition matrix and assumptions regarding the revenues, costs, correlations of the drug-development process—summarized in Tables 1 and 2—we perform two simulations depicted in Figure 4, labeled “Simulation A” and “Simulation B”. Simulation A corresponds to early-stage investments that begin in the preclinical phase and are sold when they transition to Phase II. Simulation B corresponds to later-stage investments in which compounds are acquired in Phase II and sold when they are FDA-approved. This division acknowledges the significant scientific and
business differences between early-stage investments, which are typically the domain of VCs, and later-stage development typically undertaken by large biotech or pharma companies that license compounds pre-developed by smaller biotech companies and finance their development until discontinuation or FDA approval. By structuring the simulation in two stages, we are implicitly allowing different sets of investors to participate during different phases of the drug-development process. This structure permits the maturities of the bonds to be much shorter than would be the case if compounds were funded by the same investors throughout the full cycle from preclinical development to FDA approval, which can often exceed a decade. Full-cycle simulations can also be performed within our framework. Taken together, the two simulations performed in this paper provide a compelling case for megafund financing throughout the entire drug-development cycle.

The simulation experiments are done in pairs, each pair consisting of a traditional all-equity fund—similar to a VC or mutual fund—versus a matching RBO securitization structure with a senior tranche, a junior tranche, and an equity tranche, where the size of the equity investment is the same in both (we use three tranches only for expositional simplicity; in practice, more tranches could be offered). Unlike the illustrative example above in which we assumed that individual drug-development programs yielded uncorrelated cashflows, our simulations do impose a more realistic 20% positive correlation between the valuations of all pairs of compounds to capture the potential for the clustering of negative outcomes in any given period (see SI for details).

Table 3 presents the results of a megafund with $5 billion of initial capital invested over seven and a half years in a target portfolio of 100 preclinical and 100 Phase-I compounds, with a
$1.25 billion senior tranche, a single $1.25 billion junior tranche, and a $2.5 billion equity tranche, implying a modest leverage ratio of 2-to-1 for the SPV. In a simulation consisting of 500,000 independent sample paths, an average of 102 compounds reached the goal of entering Phase II. Since there is a non-zero probability in our transition matrix $P$ of transitioning from Phase I to Phase III, the simulation also generated a small number of compounds that reached Phase III, NDA, and APP before the end of the life of the fund.

Table 3 shows that the megafund is almost always profitable. The senior-tranche RBO investors received an annual yield of 5% and were repaid in full 99.9% of the time, which is comparable to historical default rates of the highest-rated bonds according to Moody’s and Standard & Poor’s, the two largest bond-rating agencies in the industry. Junior-tranche RBO investors were paid an annual yield of 8% and repaid in full 99.1% of the time. Finally, equity-tranche investors received an average annual return of 8.9% per year, and in over a third of the simulated sample paths their average annual return exceeded 15%. While such returns may not be sufficiently attractive to traditional VC investors, large institutional investors such as pension funds, insurance companies, and endowments are likely to show more interest. Recall that unlike VC funds and all-equity structures where the possibility of substantial or total loss can be nontrivial, the megafund structure offers both debt and equity—risk-seeking investors can purchase the latter and more conservative investors can participate via the former. Since there are substantially larger pools of conservative investment capital, RBOs allow the biopharma industry to greatly expand its drug-development efforts by tapping into this tremendous asset base. In fact, certain types of financial institutions may find RBOs especially attractive either because they serve as natural hedges to ex-
isting business risks, e.g., annuity providers (whose revenues decline when people live longer), or because their corporate mandate is to support socially relevant activities but precludes them from investing in equity (in which case they can now invest a portion of their endowment’s assets in RBOs rather than just awarding grants from the annual interest on those assets).

The higher risk of the equity tranche is accompanied by the benefits of leverage provided by the bond issue, which allows the SPV to invest in a larger and more diversified portfolio of assets. This effect can be quantified by comparing the results of the equity-and-debt case with the all-equity simulation, also reported in Table 3, in which the SPV contains the same amount of equity capital ($2.5B) but no debt. In the all-equity simulation, the megafund invests in 50 preclinical and 50 Phase-I drugs, successfully carrying 52 to Phase II and generating an expected annualized return of 7.2%. The fact that this is lower than the 8.9% return in the RBOs case is explained by the lower risk of the less-leveraged portfolio (note that the probability of a negative return is 17% in the all-equity case and 20% in the RBOs case).

In Simulation B, compounds are acquired in Phase II and each can transition to its next development phase, be discontinued, or sold. Any compounds that are approved for the market are automatically sold. Table 3 presents the results of 500,000 independent simulated sample paths of a megafund with $15 billion of initial capital invested over seven and a half years in a target portfolio of 100 Phase-II compounds. The capital structure consists of a $6 billion senior tranche (with 5% yield as in Simulation A), a $3 billion junior tranche (with 8% yield), and a $6 billion equity tranche, implying a 2.5-to-1 leverage ratio. On average, the simulation yielded
just under 8 compounds approved for sale and over 21 compounds advanced to Phase III or NDA (because they did not have time during the life of the fund to reach market approval or were sold to finance principal and interest payments to bondholders). The investment performance of this SPV is even more attractive than the early-stage simulation. Senior-tranche RBO investors were repaid in full 99.9% of the time, junior-tranche investors were repaid in full 99.4% of the time, and equity-tranche investors received an average annual return of 11.4%, which compares favorably with the results offered by a comparable equity-only fund. In fact, an equity-only fund with the same equity capital ($6B) would finance the development of 40 Phase-II drugs, with only 6 advancing to Phase III or NDA, 5 to market, and offering investors an expected annualized return of only 7.2%. Securitization creates more value by engaging a larger and more diversified pool of investors, allowing larger pools of drugs to be developed.

Will rates of return of 8.9% to 11.4% for equity and 5% to 8% for debt attract capital of $5 to $15 billion as we have assumed in these simulations? The answer depends on the type of investor. Such returns may be of little interest to the private-equity and venture capital community given the smaller-scale, higher-risk/higher-return nature of their typical investments. However, for more conservative and larger institutional investors such as pension funds, insurance companies, money-market and mutual funds, endowments, foundations, and trusts, our simulated returns may be more compelling. To see why, consider the fact that the median rates of investment return of public pension fund assets over the 5-, 10-, 20-, and 25-year periods ending June 30, 2011 were 4.7%, 5.7%, 8.5%, and 8.5% respectively.\(^{25}\) Moreover, in a 2010 survey of 126 public pension funds with assets totalling $2.3 trillion, the median target investment return was 8.0%.\(^{26}\) This number
represents more than just a survey response—it is incorporated as an actuarial assumption that affects a pension fund’s investment and pension-benefit decisions, hence a target return of 8.0% is a relevant hurdle rate for such institutions. Of course, institutional investment decisions also depend on other characteristics besides return potential such as risk and correlation to broadly diversified stock and bond portfolios (which are the vast majority of these institutions’ holdings). These considerations are precisely the motivation for offering multiple tranches, each with a different risk/reward profile.

Discussion

The megafund structure can be extended and customized in many respects to suit a variety of business configurations and investor needs, e.g., more tranches, staggered debt maturities, payments that are contingent upon certain research milestones, and so on. However, despite these potential enhancements and the promising simulation results, the many embedded parameters and assumptions suggest that more research is needed before this approach becomes truly operational. In some cases, our assumptions are conservative. For example, our assumed 8% and 5% coupon rates for junior- and senior-tranche RBOs are higher than those required by today’s investors; we assumed a probability-adjusted cost of developing a compound of over $1.2 billion; and we ignored potentially significant synergies and cost savings likely to accrue to a large entity involved with multiple anticancer-therapy teams. However, in other respects, our assumptions are quite ambitious. We have implicitly assumed that there is a sufficient supply of anticancer compounds to meet the demands of our megafund; that several billion dollars of capital can be deployed in

22
high-quality research programs over a short startup period; that there is no shortage of talented researchers, engineers, and entrepreneurs who will staff the various teams needed to develop these compounds; that compounds that are not discontinued can be sold within one year at some random price drawn from a lognormal distribution; and that a very large number of compounds can be developed simultaneously and efficiently by multiple teams working in parallel.

But one of the most speculative assumptions underlying our simulations is that historical data can be used to extrapolate the future returns to drug-development efforts. New trends and nonstationarities in the stochastic process of biomedical R&D may reduce the accuracy of such extrapolations. For example, the fantasy of personalized medicine is fast becoming a reality through advances in pharmacogenomics and the identification of genetic and molecular biomarkers for various types of cancer.\textsuperscript{27} This recent innovation has had dramatic impact on the biopharma industry, creating smaller and less correlated biotech niches but also inducing greater correlation among big pharma companies that are targeting the same molecular pathways.\textsuperscript{28} The parameters of our simulation are clearly affected by such innovations, hence these and other context-specific considerations should be incorporated in any live application of our approach.

However, the inability to accurately predict research outcomes does not imply an inability of investors to assess the financial risks of and commit capital to a diversified portfolio of such outcomes. With sufficient scale, time, and expertise, biomedical megafunds may well yield attractive investment opportunities to a much broader universe of investors than those who currently invest in the biopharma industry. While securitization addresses scale and time, obtaining the nec-
cessary expertise requires an unprecedented level of collaboration between academia and industry, and among doctors, scientists, and engineers,\textsuperscript{29} including financial engineers. By incorporating more specific knowledge about industry trends, transformative scientific discoveries, and potential interactions between various drug-development programs into a megafund’s portfolio construction process, investment performance can be improved substantially. More importantly, domain-specific expertise can provide more accurate risk assessments of a megafund’s holdings, reducing the element of surprise for investors, portfolio managers, and researchers.

The fact that such collaboration does not yet exist may be a symptom of a deeper divide between academia and the biopharma industry: a cultural gap between scientific research and commercial enterprise. In a comprehensive study of the business of science—with particular emphasis on biopharma—Pisano (2010) provides an eloquent summary of this gap:\textsuperscript{2}

Science is a world focused on “first principles” and methods; in contrast, business worries about commercially feasible products and processes. Science is inhabited by academics; the manager, the industrial scientists, and the engineer dominate business. Both science and business are intensely competitive worlds but their markets and currency are distinct. In science, score is kept by peer review and grant givers, and measured ultimately by reputation; in business, score is kept by capital markets and measured by profitability. Publication is synonymous with science, secrecy synonymous with business.
While the RBO structure with long-term debt relieves some of the exigencies of shorter-term financing, it does not address the fundamental conflict between science and business. However, the new structure of a biomedical megafund presents an opportunity to re-engineer the business of science by combining into a single portfolio all of the necessary expertise to take on the great challenges of the industry.

By forming a large entity dedicated to eradicating disease, the biomedical megafund can significantly increase public awareness for both the burden of disease and the potential for its cures, allowing the fund to gather proportionally greater resources to achieve its mandate. These resources involve more than just capital, long horizons, and financial diversification. They also include: research synergies, efficiency gains, and greater collective intelligence among multiple R&D teams (who would otherwise be prevented from exchanging ideas if they worked for unrelated competing companies); centralized management of clinical trials and shared information about their outcomes (especially negative results, which are currently not reported anywhere); complementary educational synergies (e.g., facilitating a larger pipeline of M.D./Ph.D.s); stronger political support due to higher visibility among voters (e.g., government sponsorship); and greater drawing power for hiring leading experts.

This last feature of a biomedical megafund may be the most effective way to bridge the cultural gap between scientists and business executives. The most talented researchers may not be motivated by financial gain. However, an opportunity to join an elite team of like-minded individuals devoted to a worthy humanitarian challenge—with more resources at their disposal than
any other biopharma company in the history of the industry, supported by the very best engineers and clinicians, with more patient investors than the longest-horizon VC fund, and an organization solely focused on reducing the burden of disease—may be considerably more compelling. And with current challenges such as cancer, heart disease, dementia, Alzheimer’s disease, diabetes, obesity, malaria, and influenza, there is no shortage of projects with great social significance to support several biomedical megafunds. Large-scale diversified drug-development efforts facilitated by megafunds not only increase the likelihood of success, but also increase the economic value of these enterprises to all stakeholders. With sufficient scale, it becomes possible to do well by doing good.

Because ethical and humanitarian considerations affect the biopharma industry more directly than others, financial innovations that explicitly address such concerns may also be worth exploring. For example, if a cure for a lethal type of cancer is developed, should its price be whatever the market will bear? Government intervention would almost surely impose price limits because of the ethical considerations surrounding life-or-death choices linked to economic profit. This inevitable consequence of the political economy of public health suggests that pure profit-maximizing behavior in the life sciences industry may not be sustainable. One approach to addressing this issue is to incorporate broader social objectives directly into the capital structure of the megafund. For example, beyond a certain threshold of profitability, a portion of the megafund’s excess profits might be used to subsidize drugs for those least able to afford them. The immediate effect of such corporate policies would likely be reduced upside potential for equity-tranche investors; no doubt, the market price of the RBOs’ equity tranche would adjust swiftly to reflect
these new policies. However, if such policies lead to a larger and more sustainable enterprise and broaden the appeal of the securities issued under such terms to a wider investor population, the ultimate impact on the amount of capital raised and the megafund’s likelihood of success may, in fact, be positive. This possibility bears further investigation.

The megafund structure also has some potential pitfalls, including the possibility of too much centralization and control which can sometimes stifle the creativity and independence of the individual projects being funded. The corporate governance structure of the megafund must be carefully crafted to promote collective intelligence while maintaining focus on the overall purpose of easing the burden of disease. Also, successful megafund financing for biomedical applications may create new challenges of its own, such as the recent systemic failures of the mortgage-backed securities markets. Rules regarding sales practices, disclosure requirements, permissible corporate governance structures, and suitability criteria for investors must be imposed and strictly enforced to ensure good practices are observed in these important markets. These issues are beyond the scope of this paper, but our analysis serves as a proof-of-concept of what may be possible if we can resolve them.

Conclusion

Cancer is just one of a growing number of large-scale challenges confronting modern society that can only be addressed through the sustained collaboration of thousands of highly skilled, dedicated, and independent individuals over many years. Financial engineering methods such as portfolio
theory and securitization facilitate such complex collaborations by providing appropriate financial incentives to all stakeholders. Although altruism and charitable giving are important elements in responding to these challenges, we cannot rely solely on these motivations given the scale of the problems to be solved. By structuring biomedical research funding in an RBO format, incentives to reduce the burden of disease are distributed to a much broader community of stakeholders. As a result, significantly greater resources can be marshalled to take on such challenges which, in turn, will attract leading experts to join the effort, instilling even more confidence among investors, and so on. Such a “virtuous cycle” can greatly magnify a megafund’s likelihood of success. Several important government initiatives are already underway for speeding up translational medical research such as the NIH’s National Center for Advancing Translational Sciences and the Israeli Life Sciences Fund. But with budgets of only $575 and $200 million, respectively, these efforts will eventually also require substantial private-sector funding—megafunds may be one solution.

Our proposed application of securitization may be untested, but the techniques are not. They have been, and are still, used extensively in the financial industry. Some of these uses involve mortgage-related securities that played a central role in the recent financial crisis, which has created a backlash of skepticism and distrust among certain investors and issuers. Rather than shying away from such techniques because of the crisis, a more measured response may be to acknowledge their strengths, address their weaknesses, and use them wisely for the most pressing social priorities. Despite the lack of consensus regarding the ultimate causes of the financial crisis, its magnitude provides compelling evidence that with the proper incentives and financial structure, tremendous amounts of capital can be gathered in a relatively short period of time. If deployed responsibly and
effectively, such capital could transform a number of industries currently facing similar difficulties.

Proposing to raise billions of dollars for biomedical research in the current economic climate may seem ill-timed and naïve. However, today’s low-interest-rate environment is, in fact, ideal for issuing long-term debt, and investors around the globe are desperately seeking new investment opportunities that are less correlated with traditional asset classes. More importantly, the cost in terms of burden of disease—as measured by the more than half a million people expected to die of cancer this year in the U.S. alone or the $263 billion in estimated economic impact—must be balanced against the risk of failure. Similar trade-offs exist for other grand challenges of this century such as flu pandemics, climate change, and the energy crisis. Instead of asking whether we can afford to invest billions more at this time, perhaps we should be asking whether we can afford to wait.


23. DiMasi, J. & Grabowski, H. The cost of biopharmaceutical r&d: Is biotech different? *Man-

24. Standard & Poor’s Rating Services. 2010 annual global corporate default study and rating
transitions (2011).

Association of State Retirement Administrators (2011).

Survey (2011).

27. MacConaill, L., van Hummelen, P., Meyerson, M. & Hahn, W. Clinical implementation of
comprehensive strategies to characterize cancer genomes: Opportunities and challenges. *Can-


32. Frangioni, J. V. You can’t give it away in the USA, or can you? Tech. Rep., Center for Molecular Imaging, Beth Israel Deaconness Medical Center (2012).

**Supplementary Information** This information is provided in the “Supplementary Information” section below, and will be made available online in the published version of this manuscript.

**Acknowledgements** We thank Jim Broderick for his advice and help in understanding the challenges of developing and financing new drugs, and for motivating our interest in this area in discussions with A.W.L. in 2007. We also thank Janice Reichert and the Center for the Study of Drug Development (Tufts University School of Medicine) for generously sharing their data with us and for her guidance and support in accessing this data, and Lisa Natanson of Deloitte Recap LLC for giving us access to Recap’s DEAL builder™ and DEVELOPMENT optimizer™ online tools which contained detailed information on the economics of licensing deals and clinical trials of oncology compounds. Lloyd Han, James Noraky, Ashutosh Singhal, and Christopher Wilfong provided excellent research assistance throughout the entire project, and we also acknowledge Allister Bernard, Helen Hway Chen, Ming Jack Chen, Rumela Das, and Roman Garcia for research support during various phases of the project. Finally we thank Jim Broderick, Lewis Cantley, Shirish Chinchalkar, John Cox, Ora Dar, Ashish Das, John Frangioni, Jacob Goldfield, Tom Kalil, Eugene Kandel, Michael Kanef, Andrew Kimball, Doug Jamison, Eric Lander, Monique Mansoura, Albert Metz, Bob Merton, Fiona Murray, Krishna Murthi, Duane Roth, Melissa Stevens, Kailash Swarna, Alastair Wood, Glenn Yago, and participants at the Financial Innovations Lab Workshop hosted by the Milken Institute and FasterCures in July 2011 for helpful comments and discussion. The views and opinions expressed in this article are those of the authors only and do not represent the views and opinions of MIT, AlphaSimplex, Moody’s Corporation, any of their affiliates or employees, or any of the individuals acknowledged above. Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged.

**Author Contributions** All authors contributed equally to this research. A.W.L. first developed the idea for securitizing biomedical research after conversations with Jim Broderick in March 2007 about a portfolio approach to biomedical innovation. A.W.L. assembled key members of the project team, provided funding through the MIT Laboratory for Financial Engineering, and was responsible for overall project management. JM.F. was responsible for coordinating all aspects of the project, including directing research assistants, obtaining and processing all input data, calibrating the simulation parameters, running the simulations, and preparing the initial draft of the manuscript, with input and oversight from A.W.L. and R.M.S. R.M.S.
developed the analytic framework for modeling the portfolio of drug compounds. R.M.S. and Lloyd Han developed the R code with assistance from James Noraky and JM.F, and input from A.W.L. and Ashutosh Singhal. Allister Bernard converted the R code to Matlab. A.W.L. and JM.F. validated the final version of the Matlab code. R.M.S. also prepared the description of the simulation results, which was reviewed and revised by JM.F. and A.W.L. A.W.L. constructed the illustrative portfolio example and prepared the final draft of the manuscript, with input and revisions from JM.F. and R.M.S.

**Competing Interests** In addition to his MIT faculty position, A.W.L. has the following affiliations: Research Associate, National Bureau of Economic Research; Chief Investment Strategist, AlphaSimplex Group; consultant, Office of Financial Research; member, Moody’s Advisory and Academic Research Committee; member, Financial Advisory Roundtable, Federal Reserve Bank of New York; member, Economic Advisory Committee, FINRA; member, Board of Overseers, Beth Israel Deaconess Medical Center.

In addition to his MIT LFE position, R.M.S. is managing director at Moody’s Corporation.

No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No direct funding was received for this study; general research support was provided by the MIT Laboratory for Financial Engineering and its sponsors. The authors were personally salaried by their institutions during the period of writing (though no specific salary was set aside or given for the writing of this paper).

**Correspondence** Correspondence regarding the manuscript should be addressed to Andrew W. Lo, MIT Sloan School of Management, 100 Main Street, E62–618, Cambridge, MA 02142–1347, USA (email: alo@mit.edu). Correspondence regarding access to the megafund simulation software should be addressed to Jose-Maria Fernandez, MIT Laboratory for Financial Engineering, 100 Main Street, E62–611, Cambridge, MA 02142–1347, USA (email: jose-maria.fernandez@sloan.mit.edu)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Preclinical</th>
<th>Phase I</th>
<th>Phase II</th>
<th>Phase III</th>
<th>NDA</th>
<th>Approved</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>16</td>
<td>30</td>
<td>82</td>
<td>425</td>
<td>1,515</td>
<td>1,870</td>
</tr>
<tr>
<td><strong>Max</strong></td>
<td>100</td>
<td>250</td>
<td>500</td>
<td>1,000</td>
<td>2,500</td>
<td>5,000</td>
</tr>
<tr>
<td><strong>µ</strong></td>
<td>2.4</td>
<td>3.0</td>
<td>4.0</td>
<td>5.8</td>
<td>7.3</td>
<td>7.2</td>
</tr>
<tr>
<td><strong>σ</strong></td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td><strong>ρ</strong></td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Investment Assumptions (in $millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upfront</td>
</tr>
<tr>
<td>Milestone</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Development Cost Assumptions (in $millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean exp cost</td>
</tr>
<tr>
<td>SD cost/phase</td>
</tr>
<tr>
<td>Max cost/phase</td>
</tr>
<tr>
<td><strong>µ</strong></td>
</tr>
<tr>
<td><strong>σ</strong></td>
</tr>
</tbody>
</table>

Table 1: Summary of valuation and cost assumptions for the biomedical megafund simulation. The means and standard deviations of the lognormal distribution of costs and valuations were calibrated based on published studies and public databases; details are provided in SI.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Assumed Value in Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time:</strong></td>
<td></td>
</tr>
<tr>
<td>Tenor of the RBO</td>
<td>7.5 years (15 semesters)</td>
</tr>
<tr>
<td>Time to deploy capital</td>
<td>1 semester</td>
</tr>
<tr>
<td>Time to sell each compound</td>
<td>2 semesters</td>
</tr>
<tr>
<td><strong>Capital Structure:</strong></td>
<td></td>
</tr>
<tr>
<td>Total amount of capital</td>
<td>Between $5 and $15 billion</td>
</tr>
<tr>
<td>Tranches</td>
<td>Senior bond, junior bond, equity</td>
</tr>
<tr>
<td>Leverage</td>
<td>2 or 2.5 times</td>
</tr>
<tr>
<td>Bond annual yield</td>
<td>Senior bond 5%, junior bond 8%</td>
</tr>
<tr>
<td>Redemption senior bond</td>
<td>Equal semiannual installments from semester 5 to 8</td>
</tr>
<tr>
<td>Redemption junior bond</td>
<td>Equal semiannual installments from semester 9 to 12</td>
</tr>
<tr>
<td>Cash out equity</td>
<td>Period 17</td>
</tr>
<tr>
<td><strong>Investor Protection Rules:</strong></td>
<td></td>
</tr>
<tr>
<td>Interest coverage test</td>
<td>Senior debt (2), junior debt (3 or 3.5)</td>
</tr>
<tr>
<td>Cash reserved at start</td>
<td>To cover for 2 periods of interest and expected drug development costs</td>
</tr>
<tr>
<td><strong>Other:</strong></td>
<td></td>
</tr>
<tr>
<td>Number of compounds per fund</td>
<td>Between 40 and 200</td>
</tr>
<tr>
<td>Equity ownership of each asset</td>
<td>85%</td>
</tr>
<tr>
<td>RBO service fee</td>
<td>0.5% per year of total assets under management</td>
</tr>
<tr>
<td>Return on excess cash</td>
<td>1% per year</td>
</tr>
</tbody>
</table>

Table 2: Additional parameters of the biomedical megafund simulation (see SI for details).
<table>
<thead>
<tr>
<th>Variable or Summary Statistic</th>
<th>Simulation A</th>
<th></th>
<th>Simulation B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All-Equity</td>
<td>RBOs</td>
<td>All-Equity</td>
<td>RBOs</td>
</tr>
<tr>
<td><strong>Number of Compounds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preclinical</td>
<td>50</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Phase I</td>
<td>50</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Phase II</td>
<td>-</td>
<td>-</td>
<td>40</td>
<td>100</td>
</tr>
<tr>
<td>Phase III</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Research Impact</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of compounds to reach Phase II</td>
<td>52.0</td>
<td>101.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of compounds sold in Phase III and NDA</td>
<td>2.1</td>
<td>2.3</td>
<td>6.0</td>
<td>21.3</td>
</tr>
<tr>
<td>Number of compounds sold once APP</td>
<td>0.6</td>
<td>1.0</td>
<td>5.1</td>
<td>7.6</td>
</tr>
<tr>
<td><strong>Liabilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital ($ million)</td>
<td>2,500</td>
<td>5,000</td>
<td>6,000</td>
<td>15,000</td>
</tr>
<tr>
<td>Senior Tranche ($ million)</td>
<td>-</td>
<td>1,250</td>
<td>-</td>
<td>6,000</td>
</tr>
<tr>
<td>Junior Tranche ($ million)</td>
<td>-</td>
<td>1,250</td>
<td>-</td>
<td>3,000</td>
</tr>
<tr>
<td>Equity Tranche ($ million)</td>
<td>2,500</td>
<td>2,500</td>
<td>6,000</td>
<td>6,000</td>
</tr>
<tr>
<td><strong>Equity Tranche Performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average annualized ROE</td>
<td>7.2%</td>
<td>8.9%</td>
<td>7.2%</td>
<td>11.4%</td>
</tr>
<tr>
<td>P (ROE&lt;0)</td>
<td>17%</td>
<td>20%</td>
<td>16%</td>
<td>10%</td>
</tr>
<tr>
<td>P (ROE&gt;5%)</td>
<td>61%</td>
<td>68%</td>
<td>63%</td>
<td>79%</td>
</tr>
<tr>
<td>P (ROE&gt;15%)</td>
<td>15%</td>
<td>35%</td>
<td>14%</td>
<td>40%</td>
</tr>
<tr>
<td><strong>Debt Tranches Performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior Tranche: Def. Pr. , E.Loss (bp)</td>
<td>-</td>
<td>1 , &lt; 1</td>
<td>-</td>
<td>6 , &lt; 1</td>
</tr>
<tr>
<td>Junior Tranche: Def. Pr. , E. Loss (bp)</td>
<td>-</td>
<td>87 , 27</td>
<td>-</td>
<td>60 , 30</td>
</tr>
</tbody>
</table>

Table 3: Performance summary statistics of the biomedical megafund simulations. “bp” denotes units of basis points or 0.01%.
Figure 1: Timeline of cashflows for illustrative example of a typical drug-development program in which out-of-pocket costs with present value $200MM at date 0 generates annual net income of $2B in years 11 through 20, which has a present value of $12.3B at date 10 (based on a 10% cost of capital).

Figure 2: Schematic of the waterfall of cashflows for a typical RBO securitization.
Figure 3: Business structure of a biomedical megafund SPV. Funds are raised from retail or institutional investors (1) through the capital markets issuance (2) of various types of debt and equity. These funds are invested in molecules being developed to cure cancer (3). Some funds are reserved to pay for later clinical development costs and, if required, to cover the first few periods of coupon payments. The portfolio of drugs is developed over time (4). At any time a compound can be discontinued, move to the next phase or even two phases ahead based on the results of the trials done. It is also possible that compounds are sold prior to their FDA approval for marketing if it is necessary to monetize them to cover some of the fund interest or principal payments. Any compound that is approved for marketing as a new drug is sold to a biopharma company. At the end of the life of the fund, all remaining compounds in the portfolio are sold (5). After bondholders are paid back (6), the residual cash is used to pay back the equityholders (7).
Figure 4: Simulating two distinct business stages of a biomedical megafund.
Can Financial Engineering Cure Cancer?:
A New Approach for Funding Large-Scale
Biomedical Innovation

SUPPLEMENTARY INFORMATION

Jose-Maria Fernandez\textsuperscript{1}, Roger M. Stein\textsuperscript{1,2}, Andrew W. Lo\textsuperscript{1,3,4}

\textsuperscript{1}MIT Sloan School of Management and Laboratory for Financial Engineering,
100 Main Street, E62–618, Cambridge, MA 02142, USA
\textsuperscript{2}Moody’s Corporation, 7 World Trade Center,
New York, NY 10007, USA
\textsuperscript{3}MIT CSAIL and EECS, 32 Vassar Street, 32–G608,
Cambridge, MA 02139, USA
\textsuperscript{4}AlphaSimplex Group, LLC, One Cambridge Center,
Cambridge, MA 02142, USA

1 Illustrative Example

Consider a single Bernoulli trial $I_i$ with probability $p$ of success ($I_i = 1$) and probability $1 - p$ of
failure ($I_i = 0$). Then the probability of at least one success in $n$ independently and identically
distributed (IID) trials is:

$$\Pr\left(\sum_{i=1}^{n} I_i \geq 1\right) = 1 - \Pr\left(\sum_{i=1}^{n} I_i = 0\right) = 1 - (1 - p)^n. \quad (1)$$

For $p = 0.05$ and $n = 150$, this probability is $1 - 0.95^{150} = 0.9995$.

Assume that the drug-development process takes 10 years, and a success implies annual net
income of $X$ per year from years 11 to 20 (see Figure 1). Then the date-10 present value $Y_{10}$ of
this stream of cashflows is:

$$Y_{10} = \frac{X}{r} \left(1 - \frac{1}{(1 + r)^{10}}\right) \quad (2)$$

where $r$ is the cost of capital associated with cashflows $\{X\}$. For $X = $2 billion and $r = 0.10$,
$Y_{10} = $12.3 billion.
Figure 1: Timeline of cashflows for illustrative example of a typical drug development program in which out-of-pocket costs with present value $C_0$ at date 0 generates annual net income of $X$ in years 11 through 20, which has a present value of $Y_{10}$ at date 10.

If the date-0 present value of all out-of-pocket costs required to generate $X$ is $C_0$, the total investment return $R_s$ between date 0 and date 10 for a successful outcome is:

$$R_s = \frac{Y_{10}}{C_0} - 1 = \frac{X}{rC_0} \left( 1 - \frac{1}{(1+r)^{10}} \right) - 1.$$  \hfill (3)

The annualized rate of return over this period is:

$$\text{Annualized Rate of Return} = (1 + R_s)^{1/10} - 1.$$ \hfill (4)

If $C_0 = $200 million, the annualized rate of return of a success is 51%.

The rate of return $R$ for this project can then be expressed as the following Bernoulli random variable:

$$R = \begin{cases} R_s & \text{with probability } p \\ -1 & \text{with probability } 1 - p \end{cases}$$ \hfill (5)

where $R_s$ is defined in (3). The mean and standard deviation of $R$ then follow directly:

$$E[R] = p(1 + R_s) - 1,$$

$$\text{SD}[R] = \sqrt{\text{Var}[R]} = (1 + R_s) \sqrt{p(1-p)}.$$ \hfill (6,7)

The annualized expected return and standard deviation of return during dates 0 to 10 are then:

$$\text{Annualized Expected Return} = (1 + E[R])^{1/10} - 1,$$

$$\text{Annualized Standard Deviation of Return} = \frac{\text{SD}[R]}{\sqrt{10}}.$$ \hfill (8,9)
Given the parameters assumed so far, the annual expected return and standard deviation are 11.9% and 423.5%, respectively.

For a portfolio of \( n \) identical projects that are statistically independent, the return \( R_p \) and its first two moments are:

\[
R_p = \frac{Y_{10} \sum_{i=1}^{n} I_i}{nC_0} - 1 = \frac{(1 + R_s) \sum_{i=1}^{n} I_i}{n} - 1 \tag{10}
\]

\[
E[R_p] = p(1 + R_s) - 1 \tag{11}
\]

\[
SD[R_p] = \sqrt{\frac{(1 + R_s)^2}{n} p(1-p)} = (1 + R_s)\sqrt{\frac{p(1-p)}{n}}. \tag{12}
\]

As (11)–(12) demonstrate, the expected return of the portfolio is invariant to the number of programs, but the risk of the portfolio (as measured by standard deviation) declines as the number of projects increases at a rate of \( 1/\sqrt{n} \). The annualized values of the expected return and standard deviation are given by (8)–(9) as before. For \( n = 150 \), the return standard deviation is 34.6%.

To compute the default probability of a debt-financed portfolio of projects with total asset return \( R_p \), we require the probability distribution of \( \sum_{i=1}^{n} I_i \), which is a binomial random variable with distribution function:

\[
\Pr\left( \sum_{i=1}^{n} I_i \leq k \right) = \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j}. \tag{13}
\]

The default probability of a 10-year bond at date 0 which pays no coupons and promises to pay \( F \) upon maturity at the end of year 10 is then:

\[
\Pr\left( Y_{10} \sum_{i=1}^{n} I_i < F \right) = \Pr\left( \sum_{i=1}^{n} I_i < F/Y_{10} \right) = \sum_{j=0}^{\lceil F/Y_{10} - 1 \rceil} \binom{n}{j} p^j (1-p)^{n-j} \tag{14}
\]

where \( \lceil F/Y_{10} - 1 \rceil \) denotes the smallest integer greater than or equal to \( F/Y_{10} - 1 \) (note that \( \lceil F/Y_{10} - 1 \rceil + 1 \) is the minimum number of successful projects needed to repay the debt \( F \), and we assume that \( F \) satisfies the inequality \( 0 < F/Y_{10} \leq n \). For \( n = 150 \) and \( p = 0.05 \), the probability of default for \( F = $24.6 \) billion is simply the probability of less than 2 successes out of 150 trials, which is 0.00405 according to (14). Note that the large magnitude of \( Y_{10} \) creates discreteness in debt capacity and default probabilities that may not exist in practice. For example, if \( F = $36.9 = 3 \times $12.3 \) billion, the default probability jumps to 0.0182 (the probability of less than 3 successes). More generally, the debt capacity \( F^* \) associated with a desired maximum
probability \( \delta \) is given implicitly by the solution to the following:

\[
\max_F \Pr \left( \sum_{i=1}^{n} I_i < F / Y_{10} \right) \leq \delta.
\]  

(15)

For expositional clarity, we have assumed that the \( n \) projects are statistically independent. In practice, even the most diverse set of translational medical programs will exhibit some pairwise dependence, reducing the diversification benefits of the portfolio and, consequently, the debt capacity \( F^* \). We incorporate such correlation explicitly in the simulations described below.

## 2 Credit Enhancement

The risk borne by investors participating in securitization transactions can be reduced using a number of protective features called credit enhancement mechanisms. Here we describe two types of credit enhancement that may be used individually or in combination.

The first involves the implementation of various types of structural features such as overcollateralization (through the tranching of the capital structure into different classes of securities) to increase the collateral support available for more senior bondholders, or cashflow redirection rules and triggers that accelerate payments to the more senior bondholders when certain test ratios are breached. We discuss an example of these ratios and rules in Section 4.

The second type of credit enhancement makes use of some form of external credit support in which a third party assumes some of the risks to bondholders (e.g., through a letter of credit or bond insurance). A particularly interesting form of external credit enhancement in our context involves various forms of governmental guarantees. There are precedents for such programs. For example, in the U.S. the government sponsored enterprises (GSEs) Fannie Mae and Freddie Mac were created to promote home ownership. The GSEs provide guarantees for the mortgages underlying the securitization transactions that participate in their programs. For a mortgage to be accepted as collateral in a pool of securitized assets, it has to conform to certain quality standards defined by the GSE. A recent example from the biomedical domain is the Israeli Life Sciences Funds, a venture capital fund of over $200 million jointly launched in 2011 by several branches of the Israeli government and the private sector. To mitigate the risks associated with biotech R&D, the government will assume some of the downside risk.

In light of these precedents, the National Cancer Institute or the National Institute for Health could consider providing some form of guarantee to biomedical megafunds whose collateral conformed to some pre-defined scientific or medical criteria. Alternatively, a private foundation might assume this role.

Credit support from such a benefactor could serve to boost investors’ interest in these securities and potentially allow the megafund to assume bigger risks in its investments (e.g., by investing in newer technologies or those with less certain outcomes or that target rarer diseases) while providing a mechanism for leveraging the capital available from the guarantor or benefactor.
3 The Drug Approval Process

The introduction of a new drug in the market is a highly regulated process. Countries typically have national agencies responsible for authorizing new compounds for sale. For the purposes of this study, we follow the process defined by the U.S. Federal Drug Administration (FDA). Every new pharmaceutical product must undergo a number of tests to ensure that it is safe and effective. The lifecycle of a new drug generally follows the path described below.

In the “Preclinical Phase” the company developing the drug tests the product in animal trials to produce evidence that there is reasonable cause and manageable risk to permit the compound to proceed to human studies, in accordance with FDA guidelines. Following this phase, the sponsoring company files an “Investigational New Drug” (IND) application. If the FDA approves the IND, the drug moves into “Phase I”, in which the drug is tested in a small number of healthy volunteers to monitor its absorption, metabolism, and toxicity in the body to get information about its safety and dosage. If the drug is determined to be too toxic or otherwise unfit, it is withdrawn at this point.

Compounds that successfully pass Phase I move into “Phase II”, where testing is done with a patient population that already has the disease targeted by the new compound. The sponsor of the trial defines a set of endpoints that exemplify the compound’s desired effectiveness and compares these endpoints with the results from the trials in diseased patients.

Upon successful completion of Phase II, the drug moves into “Phase III” in which the drug is tested in a large sample of patients to try to confirm safety and efficacy in a wider number of circumstances and subjects.

Following successful completion of these trials, the sponsor may submit a “New Drug Application” (NDA) or “New Biologics Application” (BLA) to the FDA. If the NDA or BLA is approved, the drug can be legally marketed in the U.S.

4 Simulation Design

In this section, we describe the specific assumptions and experimental design used to generate the simulated performance analysis for an oncology megafund. The goal of our simulations is not to define an optimal transaction structure or to defend a specific set of modeling assumptions. Instead, our intent is to demonstrate the feasibility of modeling a simple financial structure—using realistic economic and scientific assumptions—in which large-scale biomedical innovation yields potentially attractive investment and drug development properties. To encourage readers to experiment with the simulation, we provide the complete source code in R and Matlab under an open-source license that enables researchers to use, modify, and distribute it.
4.1 Time Units and Tenor

The time unit used in our simulations is a semester (six months). Alternative time steps are possible but we chose one semester to match the semiannual coupon payments of the bonds in the fund’s capital structure.

We assume that the life of the fund is 15 semesters (seven and a half years). The scheduled amortization for the bonds occurs in periods 5–8 for the senior bond and 9–12 for the junior bond (we use the term “period” to refer to particular semesters). Thus, the longest-dated bond issued has a tenor (time between issuance and maturity date) of six years. In the 15th period, any remaining assets that have not been either already sold or discontinued are sold and the revenues generated accrue to the equityholders. We have assumed that it takes one year to sell a compound. Consequently, equityholders would get paid at the end of period 17 of the simulation, provided there were no previous defaults that might have shortened the life of the megafund.

Funds and simulations can be easily structured for significantly longer durations. The tenor of the fund should be related to the expected time required for the largest number of compounds to reach their full economic value. Given that Simulations A and B replicate the development of compounds from Preclinical and Phase I to Phase II, and from Phase I to market approval, respectively, we obtained reasonable results for funds of about seven and a half years’ duration.

4.2 Assets

The assets in the portfolio are assumed to be new drugs being developed by biotech or pharma companies and targeted at curing some form of cancer. Under the current model design, those same companies would be responsible for developing the compounds. The megafund could act as a financing partner and a platform from which funding could be structured, subject to a set of rules that foster collaboration across projects, encourage individual and group success, and avoid moral hazard. Those rules, as well as the processes required to select and manage the assets targeted for the fund (including determining which assets to sell and in which phase of the process to sell them), will need to be defined for each new megafund, depending on its objectives and structure. Finding the right balance between the protection of investors and the development of new scientific solutions is one of the difficult tasks regarding the implementation of this model. The creation of a blue ribbon Scientific Committee supported by financial experts is a necessary condition for the success of this new type of vehicle. The megafund structure provides an opportunity to revisit the way drug development financing decisions are currently made and to explore new corporate governance structures and organizational designs.

In our experiments, the initial portfolio of assets is composed of compounds in either the Preclinical and Phase I stages (Simulation A) or in Phase II (Simulation B). Simulation A replicates the typical venture capital investment horizon that carries compounds from Preclinical or Phase I to Phase II, when a large pharmaceutical company may acquire or license the compound for later development. Simulation B replicates the subsequent biopharmaceutical investment and development of a compound from Phase II to market approval. Practitioners note that different skill sets and funding budgets are required for each of those horizons, which motivated our decision to split
the simulation in this manner. Taken together, the two simulations provide a compelling case for applying megafund financing throughout the full lifecycle of compound development.

We assume that all compounds are acquired during the first semester of the life of the fund and that no new compounds are acquired thereafter. The number of compounds in each of the simulations results from a two-step process: (1) we fix an amount of equity such that the all-equity fund and the RBO fund both have the same dollar amount of equity ($2.5 billion in Simulation A and $6 billion in Simulation B); (2) we determine the maximum number of compounds that we can expect to invest in using all the equity and debt raised in each of the funds. We acknowledge that this is a strong assumption. Deploying several billions of dollars in such a short period of time would require considerable prior due diligence work, some new form of collaboration with the current market players who have developed the expertise in making these types of investments (venture capitalists, biopharma companies, etc.), or a totally new approach to allocate capital across development drugs. While it may be more effective to deploy capital in a less abrupt fashion, i.e., acquiring new compounds throughout the life of the fund, modeling a dynamic and actively traded portfolio would create greater complexity in this simulation experiment. Therefore, for the purposes of our examples we adopt the simpler approach of acquiring the fund’s collateral upfront, as is common among current securitization transactions. The composition of the portfolios in our simulation experiments are given in Table 1.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Simulation A</th>
<th>Simulation B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity RBO</td>
<td>50 100</td>
<td>— —</td>
</tr>
<tr>
<td>Preclinical</td>
<td>50 100</td>
<td>— —</td>
</tr>
<tr>
<td>Phase I</td>
<td>— —</td>
<td>40 100</td>
</tr>
<tr>
<td>Phase II</td>
<td>— —</td>
<td>— —</td>
</tr>
<tr>
<td>Phase III</td>
<td>— —</td>
<td>— —</td>
</tr>
</tbody>
</table>

Table 1: Composition of initial portfolios of drug compounds in the simulations.

Even though all assets are acquired at date 0, not all of the cash available at date 0 is invested immediately. A cash reserve is required to finance future clinical trials whenever a compound transitions into a new clinical phase. The fund reserves as much cash as will be required (in expectation) to develop the compounds and to cover the interest payments on the notes for a certain number of semesters. In our simulation experiments, we reserve cash for two semesters’ worth of interest payments, i.e., the bonds do not amortize in the first year.

4.3 **Simulating Portfolio Dynamics**

The development of a new drug is a complex process that depends on various scientific and economic factors. Our simulation is based on the assumption that every compound can transition along a series of different predefined states of the approval process: Preclinical, Phase I, Phase II, Phase III, NDA or BLA, Market Approval, and Discontinuation. In our model Discontinuation and
Approval are absorbing states, i.e., a drug that is discontinued or approved can no longer transition into any other state. The assumption that drugs cannot be discontinued after being approved is explained by the fact that our megafund would sell all approved drugs to biopharmaceutical companies to be marketed immediately after being approved for their first indication. Also, in our data we observe a small probability that in some rare cases a compound may skip a phase—for example, conducting trials in Phase I and II simultaneously and transitioning directly to Phase III—so our simulation captures this behavior as well.

We simulate the evolution of compounds through the approval process as following a Markov process, which is a common modeling tool applied to systems that transition from state to state over time. This approach has been widely used in finance to represent various forms of credit risk, and we apply it in our context to model the drug-transition dynamics from one phase to the next.

For each state in the approval process, we estimate a vector of transition probabilities based on the analysis of data from the last 20 years of oncology drug development programs. The vector is made up of elements, \( p_{ij} \), each of which represents the probability that a compound transitions from state \( i \) to state \( j \) in the next time step (one semester in our simulations).

At every time step and for each compound that is still in the portfolio, i.e., the compound has not been discontinued or sold, we generate a random number \( u \) from a uniform distribution. The value of \( u \) is then compared to the vector of probabilities \( p_{ij} \) to determine whether a transition occurs in the next period and if so, to which state. Conceptually, this works similarly to spinning a roulette wheel where the slots represent the different states in the approval process and the size of each slot is proportional to the probability of being in that state one period later.

In the event that a transitioned compound ends up in either an Approval or Discontinuation slot, the compound is sold or dropped from the portfolio, respectively. Otherwise, it continues in the process, but now uses a new probability vector for the new state to which the compound has transitioned.

We implement simple rules to determine which compounds to sell during the life of the fund. In particular, we assume that all drugs that are approved are sold to biopharmaceutical companies who will ultimately be marketing and distributing them. In addition, compounds can be sold prior to approval to meet the interest, principal, or management fee payments. The values and other features of the sale process are described below.

### 4.4 Transition Probabilities

The transition probabilities were calculated in two ways. For compounds in Phase I or later, we used a research database that we constructed for this study. For Preclinical compounds, we refer to existing literature for the relevant probabilities and adjust them for the periodicity of our simulation.

For the compounds in the clinical development phases of our simulation, the transition probabilities were calibrated using two sources of historical data: the DEVELOPMENT optimizer database provided by Recap LLC and a dataset provided by the Center for the Study of Drug Development (CSDD), Tufts University School of Medicine. Recap’s DEVELOPMENT optimizer
database is built on curated clinical and regulatory histories for approximately 1,450 compounds entered into human clinical development in 2,467 distinct indications by a select group of more than 240 benchmarked biotechnology companies, i.e., the constituents of the Recap BioPortfolio Index™, since 1988. The histories are documented and updated daily using multiple primary, public sources of information, including but not limited to: U.S. Securities and Exchange Commission filings, U.S. and E.U. pharmaceutical regulatory documents captured and analyzed from the Food and Drug Administration (FDA) and the European Medicines Agency websites, peer-reviewed journal articles and scientific abstracts, government databases such as clinicaltrials.gov, and corporate press releases and investor information. The CSDD data were compiled from publicly available information reported by companies involved in the development of cancer drugs. The compounds targeted consisted of new molecular entities developed primarily for an anti-cancer indication for which an IND application was filed with the FDA and that entered clinical trials between January 1990 and the start of 2011. The compounds in the database were developed by biotechnology or pharmaceutical companies and were either therapeutic compounds or vaccines.

We merged the Recap and CSDD databases to yield a combined database of over 2,000 compounds. After removing duplicates and compounds for which there was not enough information about their start or transition dates, or that did not conform to the criteria defined in the paragraph above (e.g., compounds only approved for marketing outside of the U.S. or that were reformulations of existing drugs), we arrived at a final set of 733 compounds. The summary statistics for this final database are contained in Table 2.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Total</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved: (NDA)</td>
<td>38</td>
<td>5%</td>
</tr>
<tr>
<td>Discontinued (Phase I)</td>
<td>174</td>
<td>24%</td>
</tr>
<tr>
<td>Discontinued (Phase II)</td>
<td>171</td>
<td>23%</td>
</tr>
<tr>
<td>Discontinued (Phase III)</td>
<td>30</td>
<td>4%</td>
</tr>
</tbody>
</table>

Still in process as of end compilation period:

<table>
<thead>
<tr>
<th>Stage</th>
<th>Total</th>
<th>in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>In NDA</td>
<td>4</td>
<td>1%</td>
</tr>
<tr>
<td>In Phase I</td>
<td>17</td>
<td>2%</td>
</tr>
<tr>
<td>In Phase II</td>
<td>221</td>
<td>30%</td>
</tr>
<tr>
<td>In Phase III</td>
<td>76</td>
<td>10%</td>
</tr>
</tbody>
</table>

| Total   | 733   | 100% |

Table 2: Composition of the final database of 733 oncology compounds in various clinical phases (percentages do not sum to 100% due to rounding).

Using these data, we calculated the transition probabilities by first estimating a continuous-time generator matrix and then converting this to transition matrices in a discrete-time setting where the time unit is the semester.3

For compounds in the Preclinical phase, data is more difficult to collect, in part because drug development companies have little incentive to provide information about their research and
sometimes unsuccessful programs. To estimate transition probabilities for Preclinical compounds, we used statistics reported in Paul et al. (2010) and also adopted the definition of the Preclinical period used in that paper. We then scaled these long-run probabilities to arrive at the probabilities for a single semester. In doing so, we assume that compounds in this phase could only transition from Preclinical to either Discontinuation or Phase I, and that the mean time in the Preclinical phase was one year, as reported in Paul et al. (2010).

The resulting transition matrix estimate is given by:

\[
P = \begin{pmatrix}
0.5 & 0.35 & 0.00 & 0.00 & 0.00 & 0.15 \\
0.0 & 0.80 & 0.14 & 0.01 & 0.00 & 0.05 \\
0.0 & 0.00 & 0.86 & 0.06 & 0.00 & 0.08 \\
0.0 & 0.00 & 0.00 & 0.85 & 0.07 & 0.06 \\
0.0 & 0.00 & 0.00 & 0.00 & 0.56 & 0.40 \\
0.0 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \\
0.0 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \\
\end{pmatrix}
\]

The mean long-term transition probabilities (the limiting probabilities) and mean times to transition resulting from this probability matrix over the period of time covered by the study are presented in Table 3:

<table>
<thead>
<tr>
<th></th>
<th>Preclinical to Phase I</th>
<th>Phase I to Phase II</th>
<th>Phase II to Phase III</th>
<th>Phase III to NDA</th>
<th>NDA to Approval</th>
</tr>
</thead>
<tbody>
<tr>
<td>P (Transition)</td>
<td>69.0%</td>
<td>72.4%</td>
<td>45.2%</td>
<td>58.6%</td>
<td>95.2%</td>
</tr>
<tr>
<td>Avg (months in phase)</td>
<td>12.0</td>
<td>31.2</td>
<td>38.6</td>
<td>39.6</td>
<td>13.8</td>
</tr>
</tbody>
</table>

Table 3: Average transition probabilities and time per development phase.

These phase transition probabilities are comparable to those reported elsewhere in the literature for similar compounds, as shown in Table 4.4–7

<table>
<thead>
<tr>
<th>Source</th>
<th>Time Period</th>
<th>Number of Compounds</th>
<th>Preclinical to Phase I</th>
<th>Phase I to Phase II</th>
<th>Phase II to Phase III</th>
<th>Phase III to NDA</th>
<th>NDA to Approved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Megafund*</td>
<td>1990–2010</td>
<td>733</td>
<td>69.0%</td>
<td>72.4%</td>
<td>45.2%</td>
<td>58.6%</td>
<td>95.2%</td>
</tr>
<tr>
<td>Natanson*</td>
<td>1988–May 2010</td>
<td>164</td>
<td>—</td>
<td>72.6%</td>
<td>40.3%</td>
<td>66.7%</td>
<td>90.6%</td>
</tr>
<tr>
<td>Reichert et al.*</td>
<td>1990–2006</td>
<td>920</td>
<td>—</td>
<td>78.0%</td>
<td>43.0%</td>
<td>52.0%</td>
<td>89.0%</td>
</tr>
<tr>
<td>Walker et al.*</td>
<td>1995–2007</td>
<td>974</td>
<td>—</td>
<td>77.0%</td>
<td>44.0%</td>
<td>52.0%</td>
<td>—</td>
</tr>
<tr>
<td>Dimasi et al.</td>
<td>1993–2002</td>
<td>838</td>
<td>—</td>
<td>76.8%</td>
<td>59.4%</td>
<td>57.1%</td>
<td>—</td>
</tr>
<tr>
<td>Paul et al.</td>
<td>15 years</td>
<td>—</td>
<td>69.0%</td>
<td>54.0%</td>
<td>34.0%</td>
<td>70.0%</td>
<td>91.0%</td>
</tr>
</tbody>
</table>

*These probabilities are calculated only for cancer related compounds.

Table 4: Comparison of cancer compound transition probability by development phase.
4.5  Asset Valuations

In this section, we discuss the distributional model we use for simulating the values of drug compounds at the time of sale. An important feature of our model is the presence of correlation among the valuations of different compounds, which can be observed empirically to some degree and is often noted by venture capitalists and other experts in this domain. We begin by assuming that the market values of drug compounds are approximately lognormally distributed, which implies a larger number of moderately low valuations interspersed with a smaller number of large “blockbuster” valuations. In addition, we induce correlation among the valuations of compounds in the collateral portfolio. Although correlation does not affect the mean return on the overall portfolio, because of the overlay of tranching and the impact of the waterfall rules, the mean return and variability of returns on specific tranches will be affected by contemporaneous clusters of particularly high or particularly low valuations.

To induce correlation among the sale values in our model, we begin by generating correlated standard normal random variables which we then rescale and exponentiate such that the resulting (correlated) lognormal random variables have means and standard deviations that are consistent with the valuation assumptions described in Section 4.6. This mechanism may be thought of as specifying a stochastic process for the value of a compound in which there is a common unobservable factor.

Next, we specify that the value $Z_{ij}$ of the standard normal draw for the $i$th entity in the $j$th economic state of the world is the sum of two components: (1) a common (systematic) component that affects all valuations in the portfolio in economic state of the world $j$; and (2) a compound-specific (idiosyncratic) component that only affects compound $i$ in economic state of the world $j$. More formally:

$$Z_{ij} = \beta_i^S S_j + \epsilon_{ij}, \quad S_j \sim \mathcal{N}(0, 1), \quad \epsilon_{ij} \sim \mathcal{N}(0, 1 - (\beta_i^S)^2) \tag{16}$$

and $\beta_i^S$ describes the magnitude of the impact of the systematic factor $S_j$ on $Z_{ij}$. We assume that all error terms and cross terms are mutually statistically independent.

Under these assumptions, for compounds with a common value $\beta_i^S = \beta_S$, and recalling that $Z_{ij} \sim \mathcal{N}(0, 1)$, it can be shown that

$$\beta_s = \sqrt{\rho}, \quad \epsilon_{ij} = \sqrt{1 - \rho} \eta_{ij} \tag{17}$$

where $\eta_{ij} \sim \mathcal{N}(0, 1)$ is IID over $i$ and $j$, and $\rho$ is the correlation between the $Z_{ij}$ (note: $\rho$ must be greater than or equal to 0 and less than or equal to 1). This specification is similar to common one-factor credit models with correlated defaults (however, in our case the transition probability is not driven directly by the asset value—as might be the case in models of correlated default—since the transition probability is determined primarily by the drug approval process). In cases in which there is a common correlation among all compounds in the portfolio ($\beta_i^S = \beta_S$ for all $i$), the value...
of the $i$th compound in the $j$th simulation path is given as

$$X_{ij} = \exp \left( m + \left( \sqrt{\rho S_j} + \sqrt{1 - \rho \epsilon_{ij}} \right) s \right) = \exp (m + Z_{ij}s)$$

(18)

where

$$m = \ln(m_v) - \frac{1}{2} \ln \left( 1 + \frac{s^2_v}{m_v^2} \right) \equiv \text{the estimated mean of log-returns}$$

$$s = \left[ \ln \left( 1 + \left( \frac{s^2_v}{m_v^2} \right) \right) \right]^{1/2} \equiv \text{the estimated SD of log-returns}$$

(19)

and $m_v$ and $s^2_v$ are the estimated mean and variance, respectively, of the observed valuation data (or are derived from other qualitative approaches).

To avoid very large (and potentially unrealistic) simulated values for compounds at the time of sale, we also impose a maximum ($M_i$) on the value of the compound such that the final value is given as

$$V_{ij} = \min(X_{ij}, M_i)$$

(20)

The introduction of an upper bound affects the mean of the distribution. We adjusted the values of $m$ to accommodate the capping that occurs as a result of the imposition of the upper bound. This adjustment ensures that the mean of the distribution is consistent with our data. The values for $m$, $s$, and $M_i$ are presented in Table 5 for each compound phase.

In our simulations, drug compounds are sold infrequently and tend not to cluster in time except at the end of the transaction. Thus, simulations in which the (unobservable) systematic component is updated each period result in relatively little correlation among prices, even though our simulation horizons are short (i.e., on the order of 5–7 years). Furthermore, because the current structure of the portfolio is static, i.e., the portfolio of compounds is purchased at the beginning of the fund transaction and then winds down, it is natural to think of a common factor affecting the whole portfolio. Such a factor could be general economic conditions, regulatory shifts, or sweeping technological advances.

Accordingly, we assume that the valuations of compounds are lognormally distributed and governed by the dynamics in (16), and that the systematic shock occurs once at the beginning of each simulation path, such that a particular string of sales within that path will all be influenced by the shock. This induces correlation among all valuations in the portfolio, rather than just between those in which sales occur in a single period.

To implement this approach, in each path we simulate a single value for $S_j$ (in the $j$th path) and then, for each compound being sold, simulate $\epsilon_{ij}$ and calculate a valuation for compound $i$ ($V_{ij}$) as in (20). Importantly, although $S_j$ is drawn only once per trajectory, the compounds in the portfolio themselves evolve (e.g., transition, get funded, etc.) in each time step of the simulation.
4.6 Calibration of Valuation Parameters

A critical set of assumptions in our simulations involves the valuations of individual drug assets at a given stage of development. Such valuations are difficult to estimate due to significant heterogeneity across the assets with respect to a compound’s scientific merits, its commercial potential, the expertise of the managers in charge of its development, etc. Furthermore, many oncology assets have traditionally been privately held, or developed as part of a larger suite of products, and thus accurate data on individual valuations are not readily available.

To estimate the mean and variance of compound valuations for our model, we used data from Bloomberg to build a dataset of initial public offerings (IPOs) since 2000, and market valuations (as of the end of the first quarter of 2011) of publicly listed companies in the U.S. focused on developing and marketing cancer compounds and which had a market capitalization of at least $5M. For each company for which we could gather enough information, the value per approved compound was calculated by dividing the market capitalization of the company by the number of approved compounds it owned. We realize that this method may overestimate the value of compounds given that this calculation assigns a value of zero to earlier stage compounds and that the company may hold other assets not considered. Alternate methods to estimate the value of compounds such as discounted future cashflows could be applied to future developments of this model. For our simulations we decided to approximate the value of a compound using observed market valuations. The resulting mean value for a marketed compound is $1.87 billion and its standard deviation is $2.24 billion.

The valuations corresponding to the compounds in earlier development phases were calculated using a binomial-tree valuation model in which the value of each compound is estimated by taking into account the probabilities of success and failure per phase, the expected values in each case, and the time required to move from one phase to the next. The inputs for the transition probabilities and times in each phase were derived from our transition matrix. The discount rates used to calculate the discounted values per phase were 15% for the Market to NDA phase, 25% for NDA to Phase III, and 30% for the earlier phases to reflect the higher risk of early-stage projects. In addition, upper bounds for valuations were imposed to prevent the model from generating unreasonably large values. These upper bounds were chosen qualitatively based on the empirical distribution of values (see Section 4.5).

To compute the standard deviation of each of the values per phase and per compound, we used data from Bloomberg. First we calculated the standard deviation of the value of marketed compounds (the $2.24 billion cited above). Next we calculated the ratio (1.19) of the standard deviation to the mean value of marketed drugs (the $1.87 billion cited above) in our Bloomberg database. Finally, we applied this ratio to the mean values per compound and per phase to estimate each of the standard deviations corresponding to each of the phases.

The mean, standard deviation, and upper bounds were used to fit the lognormal distributions from which the value of each compound is drawn in our simulations. Table 5 shows the values of the estimated original mean and upper bounds as well as the cap-adjusted $m$ and $s$ used to estimate the value of the drugs.

Finally, as a proxy for the (normal) correlations, we use the mean pairwise correlation of the
Table 5: Parameters for valuation functions.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Cap</th>
<th>$m$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preclinical</td>
<td>16</td>
<td>100</td>
<td>2.36</td>
<td>0.939</td>
</tr>
<tr>
<td>Phase I</td>
<td>30</td>
<td>250</td>
<td>2.96</td>
<td>0.939</td>
</tr>
<tr>
<td>Phase II</td>
<td>82</td>
<td>500</td>
<td>4.00</td>
<td>0.939</td>
</tr>
<tr>
<td>Phase III</td>
<td>425</td>
<td>1000</td>
<td>5.80</td>
<td>0.939</td>
</tr>
<tr>
<td>NDA</td>
<td>1515</td>
<td>2500</td>
<td>7.35</td>
<td>0.939</td>
</tr>
<tr>
<td>Approved</td>
<td>1870</td>
<td>5000</td>
<td>7.24</td>
<td>0.939</td>
</tr>
</tbody>
</table>

equity returns on small biopharmaceutical firms, also calculated using Bloomberg data. These are related somewhat to the valuations of individual compounds since it is often the case that small firms have only a single drug that they are either researching or producing. Thus the price of the firm may be related to the value of these compounds.

In our sample, the equity correlations we estimated were on the order of 20%, which is the value we use in the simulations. This value is within the range of observed correlations among public firms in typical credit-portfolio models, and is consistent with empirical estimates of publicly traded oncology biotech firms. We expect correlations associated with very small private firms to be lower than for those of public firms, and correlations among the valuations of single compounds in our simulation to be lower still.

Note also that the correlation in returns does not translate directly into an equivalent level of correlation for realized valuations, due to the calculation of the former on returns (in normal space) and the latter on levels (in lognormal space). In general we observe that valuation levels are correlated to a lower degree than are the equity returns (i.e., the correlation among valuations is lower than 0.2), which is consistent with our expectations.

4.7 Investment Structure and Development Costs

The investment structure we assume is based on the licensing framework commonly used in the biopharmaceutical industry. In our model, during the course of the drug’s development, both upfront and periodic payments are made by the megafund to finance additional research and to compensate the developers for successful completion of key milestones (such as the completion of a phase). In addition, the megafund finances all clinical trial costs. In exchange for this funding, the megafund is granted 85% of the economic value of the compound when it is sold. The remaining 15% is assumed to be retained by the founder and management team developing the compound. The structure of the payments made by the megafund is detailed below.

In practice, upfront and milestone payments for a specific compound are derived through negotiations based on the novel features and properties of the compound, its expected value, the amount of investment required to carry the compound to the next phase(s), and the negotiating power of the parties. It is therefore difficult to define what the standard terms of any particular deal
might be. For our model, in addition to the commitment to fund the development of the drugs, we estimated the upfront payments to be 40% of the expected development costs per phase and the milestones to be 50% of the upfront payments.

We confirmed the plausibility of these investment parameters through conversations with experts, and by using data from public presentations made by practitioners and the Recap DEAL builder tool. However, recent trends seem to favor smaller upfront payments and larger milestone payments. Future research may confirm this point, which would necessitate changes in the simulation parameters.

### 4.8 Drug Development Costs

We assume that development costs per phase and per compound follow a lognormal distribution with parameters based on previous results reported in the literature. Some authors have since argued that the costs reported in these studies may be overstated, but we adopt these figures to be conservative.

For compounds in preclinical and clinical phases, Paul et al. (2010) provides estimates of the cost of development at each clinical stage based on industry benchmarking data provided by the Pharmaceutical Benchmarking Forum along with fifteen years of project level data from Lilly’s R&D portfolio. Dimasi et al. (2003) bases results on survey data collected from the commercial sponsors of 68 randomly selected approved compounds in the CSDD database, representing multiple therapeutic areas including, but not limited to, oncology.

To estimate the mean cost per phase we have chosen to use the results of Paul et al. (2010) because they are more timely and they result in more conservative expected costs. However, we omitted the submission and launch costs proposed ($40 million) which appear to include launch preparation costs that we expect to be borne by the biopharma companies acquiring these compounds post-approval (recall that under our current assumptions, once drugs are approved for their first indication, they are sold to an industry incumbent for marketing and distribution).

We further increase our cost estimates in two ways. First, since the statistics in Paul et al. (2010) are calculated in 2008 dollars, we used the U.S. GDP deflator index to inflate the numbers to 2011 dollars. Separately, we adjust for the additional cost of developing cancer compounds relative to other types of therapies. Adams and Brantner (2006) analyze the capitalized cost of

<table>
<thead>
<tr>
<th>Phase</th>
<th>Upfront Payment</th>
<th>Milestone</th>
<th>Development Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preclinical</td>
<td>2.5</td>
<td>1.3</td>
<td>Random*</td>
</tr>
<tr>
<td>Phase I</td>
<td>7.5</td>
<td>3.8</td>
<td>Random*</td>
</tr>
<tr>
<td>Phase II</td>
<td>20.0</td>
<td>10.0</td>
<td>Random*</td>
</tr>
<tr>
<td>Phase III</td>
<td>75.0</td>
<td>37.6</td>
<td>Random*</td>
</tr>
</tbody>
</table>

*See Section 4.8.

Table 6: Investment costs (in $millions).
new drug development by indication and show that the cost of developing an oncology product is 20% higher than the sample mean (across all compounds in their dataset). Accordingly, for our analysis, we adjusted the costs upward by a factor of 1.2 to reflect the higher than average cost of oncology development. Both of these adjustments yield higher costs and, therefore, more conservative profits in our simulations.

Using our mean estimates, the resulting mean out-of-pocket costs invested per compound from Preclinical to the end of Phase III is $263 million.

Paul et al. (2010) do not provide an estimate of the standard deviation of the costs per phase. For our experiments, we assumed that the variability of development costs for oncology compounds is related to that presented in Dimasi et al. (2003). We then calculated the ratio of the standard deviation to the mean cost reported in this paper (ranging from 0.70 to 0.94) and applied that ratio to the adjusted costs per phase obtained as previously explained.

In addition, we imposed a maximum cost in each phase to cap the expenses incurred per compound and per phase. The sum of the cap costs assumed per phase yields a total maximum out-of-pocket cost per compound of $690 million, which is quite conservative compared to figures contained in the literature. The resulting adjusted mean costs per phase and corresponding standard deviations are shown in Table 7.

<table>
<thead>
<tr>
<th></th>
<th>Mean cost Paul et al. (2010)</th>
<th>Mean adjusted for oncology factor (in USD2011)</th>
<th>Dimasi et al. (2003) SD/Mean ratio</th>
<th>SD per phase</th>
<th>Max cost per phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preclinical</td>
<td>5</td>
<td>6</td>
<td>0.92</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>Phase I</td>
<td>15</td>
<td>19</td>
<td>0.84</td>
<td>16</td>
<td>50</td>
</tr>
<tr>
<td>Phase II</td>
<td>40</td>
<td>50</td>
<td>0.94</td>
<td>47</td>
<td>120</td>
</tr>
<tr>
<td>Phase III</td>
<td>150</td>
<td>188</td>
<td>0.70</td>
<td>132</td>
<td>500</td>
</tr>
<tr>
<td>Total</td>
<td>263</td>
<td></td>
<td></td>
<td>690</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Asset development out-of-pocket costs.

These figures can be used to estimate the parameters $m$ and $s$ of the lognormal distribution that we use to simulate development costs, which are drawn randomly for each compound in each phase. We adjusted the values of $m$ to accommodate the capping (the maximum cost per phase) to ensure that the mean of the distribution would remain consistent with the observed data. The resulting parameters used are given in Table 8.

The out-of-pocket costs per approved compound may be estimated according to Paul et al. (2010) by calculating the number of drugs needed to obtain a single approval. Under our assumptions, eight Preclinical projects are needed to bring one new drug to market on average. Starting with eight compounds in the Preclinical phase, if we multiply the expected number of compounds that transition to each phase by the expected cost per phase we get an estimate of the out-of-pocket cost to develop a new compound of $693 million. Following Paul et al. (2010) and Dimasi et al. (2003), we capitalize these costs over time to account for the cost borne by investors to finance
### Table 8: Parameters of the lognormal distribution used to simulate development costs.

<table>
<thead>
<tr>
<th></th>
<th>Adjusted Mean Cost</th>
<th>Cap Adjusted m</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preclinical</td>
<td>6</td>
<td>1.53</td>
<td>0.79</td>
</tr>
<tr>
<td>Phase I</td>
<td>19</td>
<td>2.72</td>
<td>0.73</td>
</tr>
<tr>
<td>Phase II</td>
<td>50</td>
<td>3.65</td>
<td>0.79</td>
</tr>
<tr>
<td>Phase III</td>
<td>188</td>
<td>5.06</td>
<td>0.63</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>263</strong></td>
<td><strong>-</strong></td>
<td><strong>-</strong></td>
</tr>
</tbody>
</table>

### Table 9: Comparison of development costs of a single approved drug (from Preclinical to Approval).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-of-pocket cost</td>
<td>693</td>
<td>654</td>
<td>403</td>
<td>672 / 559</td>
<td>—</td>
</tr>
<tr>
<td>Capitalized</td>
<td>1,220</td>
<td>1,104</td>
<td>802</td>
<td>1,318 / 1,241</td>
<td>868</td>
</tr>
</tbody>
</table>

### 4.9 Capital Structure and Cashflow Waterfall

In our experiments, we assume a very simple capital structure and cashflow waterfall. Our capital structure has three tranches: a senior bond, a junior bond, and an equity tranche. A more sophisticated implementation would almost certainly take advantage of a more efficient capital structure and more involved waterfall rules.

The bonds receive semiannual coupons and are amortized in equal installments over various periods of time as presented in Table 10. The senior bonds have a maturity of 4 years and their owners receive coupon and redemption payments ahead of the junior and equity-tranche holders. The junior bonds have a maturity of 6 years and they are paid back before any cashflows accrue to the equityholders.

The securities are assumed to be quite basic (fixed-rate amortizing bonds and common equity), but a more sophisticated model might make use of less standard securities to better match investors’ preferences.

In addition to overcollateralization (which involves holding more collateral than the par value of the tranche), the structure includes an interest coverage ratio test (IC test) designed to protect bondholders. The ratio is calculated as follows:
<table>
<thead>
<tr>
<th></th>
<th>Coupon (annual)</th>
<th>Amortization Schedule (start and end semester)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior Bond</td>
<td>0.05</td>
<td>Periods 5 to 8</td>
</tr>
<tr>
<td>Junior Bond</td>
<td>0.08</td>
<td>Periods 9 to 12</td>
</tr>
<tr>
<td>Equity</td>
<td>—</td>
<td>Period 15</td>
</tr>
</tbody>
</table>

Table 10: Capital structure parameters used in simulations.

- The numerator of the IC ratio is equal to the cash reserve available plus the future expected cash inflows from the sale of compounds already in process minus the management fee (50 basis points per year) minus the interest and principal redemptions due in the current period.

- The denominator of the IC ratio is the required payments for the next $k$ periods of management fees plus interest and principal ($k$ is assumed to be 2 for Simulations A and B).

If the ratio falls below the target IC level, compounds are sold to bring the IC ratio back into compliance. The sale of assets helps ensure that the Special Purpose Vehicle (SPV) will have enough funds to pay the servicing, interest, and principal payments.

Throughout the life of the biomedical megafund, waterfall rules guide the allocation of funds. The waterfall is implemented as follows:

- At the start of each period (semester) all proceeds from any consummated compound sales are added to the current cash balance.

- Also at the start of each period, each compound is tested to see if it has transitioned to a new state. Any compounds that have transitioned into the Approved state (or to the targeted phase in the megafund, i.e., Phase II in Simulation A) are sold and the cashflow from the sale is deferred until the end of the sales cycle (we assume it takes 2 semesters to organize and execute a compound sale).

- If there is sufficient cash in the cash account, payments are made in the following order:
  - The megafund management fee is paid.
  - Interest on senior bonds is paid.
  - Scheduled principal payments on senior bonds are paid.
  - Interest on junior bonds is paid.
  - Scheduled principal payments on junior bonds are paid.

- If there is not enough money to meet these obligations, some or all of the bonds are in default and the assets are liquidated. In the event of liquidation, the cash generated by the monetization of available assets flows first to the most senior bondholders followed by the junior bondholders, and any residual amount goes to the equityholders.
If the megafund is not in default, the IC test is performed. If the IC test is failed, the cash shortfall is calculated and compounds are sold to meet the shortfall and ensure compliance with the IC test.

From the remaining cash, a portion is reserved to make the servicing, interest, and principal payments over the subsequent \( k \) periods of time (\( k \) is assumed to be 2 in the simulations).

If any surplus cash remains, it is used to finance the clinical trials of any compounds that have transitioned but have not yet received funding for their new phase, starting with those compounds that are farthest along in the approval process.

After all of the above payments are made, cash at the end of the period is calculated.

If there are no bonds left outstanding, the portfolio is liquidated and all remaining proceeds, net of administrative fees, accrue to the equityholders.

### 5 Historical Healthcare and Biotech Investment Returns

Data from the ThomsonOne database, VentureXpert (VX), indicates that over the last decade the biotech and healthcare venture capital (VC) investments have exhibited significantly lower returns than in the past. This pattern suggests that venture capital investment in this sector may be suffering from a secular downturn in returns. In VX, the biotech sector includes human therapeutic biotechnology, industrial biotechnology, and biosensors, and the medical/healthcare sector covers pharmaceutical research, therapeutics, diagnostics, and other healthcare related services. The VX database provides returns from all stages of venture investment, and reports 1-year rolling-horizon internal rates of return based on cash inflows and outflows in each year. Those returns include results from both active and liquidated funds (avoiding survivorship bias in the data), and are net of management fees and carried interests.

Figure 2 contains the 1-year IRR (the blue and green lines) and trailing 10-year IRRs (the orange lines) for the biotechnology and medical/healthcare sectors, where the 10-year IRRs are computed by compounding the 1-year IRRs over the preceding decade.

A recent paper by Booth et al. (2011)\(^{15}\) also discusses the return of life sciences venture investing over the 2000–2010 period using information from a benchmarking database from Cambridge Associates (an investment advisory firm). Focusing on returns from deals (not funds), they examine the gross (including fees) pooled mean IRR for healthcare VC investing and its subsectors, and report a return of 15% for realized deals in the healthcare sector and 7.4% if unrealized deals are included.

Table 11 compares the 2000–2010 results from VX and Booth et al. (2011) (realized and unrealized deals) with the simulations results from our model. Using parameters derived from historical data on valuations from 2000 to the first quarter of 2011, the megafund yielded an annual return of 7.2% for both Simulations A and B with an all-equity capital structure. Simulations in which the capital structure consists only of equity are closest in structure to a large venture capital fund, hence a comparison of these results to those from VX and Booth et al. (2011) seems
appropriate. The VX database shows an IRR of 2.7% for healthcare and 5.7% for biotech during 2000–2010. Since these returns are net of fees, we must add back the fees before comparing them to our simulation returns. The standard VC fees are approximately 2% per annum of assets under management plus carried interest (which we estimate to be an additional 1% per annum). This implies annual gross returns of 5.7% and 8.7%, respectively. The Booth et al. (2011) estimates of 12.8% and 7.4% already include fees. The megafund simulation yields net-of-fee returns of 7.2%, or 7.7% if we add the 0.5% service fees included in our simulation. Even though the assets involved in each of these cases do not match exactly, they are similar and, based on these comparisons, our simulation results seem consistent with recent historical experience in the biopharma industry.

<table>
<thead>
<tr>
<th>Source</th>
<th>Raw</th>
<th>Gross of Fees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated Megafund</td>
<td>7.2%</td>
<td>7.7%</td>
</tr>
<tr>
<td>VX (Healthcare)</td>
<td>2.70%</td>
<td>5.70%</td>
</tr>
<tr>
<td>VX (Biotech)</td>
<td>5.70%</td>
<td>8.70%</td>
</tr>
<tr>
<td>Booth et al. (2011) (All healthcare)</td>
<td>7.40%</td>
<td>7.40%</td>
</tr>
<tr>
<td>Booth et al. (2011) (Pharmaceuticals)</td>
<td>12.80%</td>
<td>12.80%</td>
</tr>
<tr>
<td>Booth et al. (2011) (Biotech)</td>
<td>7.60%</td>
<td>7.60%</td>
</tr>
</tbody>
</table>

Table 11: Comparison of historical returns of VC equity investment from 2000–2010 and simulated megafund returns.

Finally, it is important to note the downward trend in investment returns in the biotechnology and pharmaceutical venture capital sector over the past decade. In fact, if the year 2000—the last year the biopharma industry experienced large positive performance—is dropped from the sample, the IRRs for biotech and healthcare become negative according to VX (−0.5% for biotech and −0.7% for the medical/healthcare). This sensitivity to outliers suggests the importance of moni-
monitoring the investment performance of both sectors so as to recalibrate the simulations as needed.


**Acknowledgements**  We thank Jim Broderick for his advice and help in understanding the challenges of developing and financing new drugs, and for motivating our interest in this area in discussions with A.W.L. in 2007. We also thank Janice Reichert and the Center for the Study of Drug Development (Tufts University School of Medicine) for generously sharing their data with us and for her guidance and support in accessing this data, and Lisa Natanson of Deloitte Recap LLC for giving us access to Recap’s DEAL builder™ and DEVELOPMENT optimizer™ online tools which contained detailed information on the economics of licensing deals and clinical trials of oncology compounds. Lloyd Han, James Noraky, Ashutosh Singhal, and Christopher Wilfong provided excellent research assistance throughout the entire project, and we...
also acknowledge Allister Bernard, Helen Hway Chen, Ming Jack Chen, Rumela Das, and Roman Garcia for research support during various phases of the project. Finally we thank Jim Broderick, Lewis Cantley, Shirish Chinchalkar, John Cox, Ora Dar, Ashish Das, John Frangioni, Jacob Goldfield, Tom Kalil, Eugene Kandel, Michael Kanef, Andrew Kimball, Doug Jamison, Eric Lander, Monique Mansoura, Albert Metz, Bob Merton, Fiona Murray, Krishna Murthi, Duane Roth, Melissa Stevens, Kailash Swarna, Alastair Wood, Glenn Yago, and participants at the Financial Innovations Lab Workshop hosted by the Milken Institute and FasterCures in July 2011 for helpful comments and discussion. The views and opinions expressed in this article are those of the authors only and do not represent the views and opinions of MIT, AlphaSimplex, Moody’s Corporation, any of their affiliates or employees, or any of the individuals acknowledged above. Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged.

Author Contributions All authors contributed equally to this research. A.W.L. first developed the idea for securitizing biomedical research after conversations with Jim Broderick in March 2007 about a portfolio approach to biomedical innovation. A.W.L. assembled key members of the project team, provided funding through the MIT Laboratory for Financial Engineering, and was responsible for overall project management. JM.F. was responsible for coordinating all aspects of the project, including directing research assistants, obtaining and processing all input data, calibrating the simulation parameters, running the simulations, and preparing the initial draft of the manuscript, with input and oversight from A.W.L. and R.M.S. R.M.S. developed the analytic framework for modeling the portfolio of drug compounds. R.M.S. and Lloyd Han developed the R code with assistance from James Noraky and JM.F, and input from A.W.L. and Ashutosh Singhal. Allister Bernard converted the R code to Matlab. A.W.L. and JM.F. validated the final version of the Matlab code. R.M.S. also prepared the description of the simulation results, which was reviewed and revised by JM.F. and A.W.L. A.W.L. constructed the illustrative portfolio example and prepared the final draft of the manuscript, with input and revisions from JM.F. and R.M.S.

Competing Interests In addition to his MIT faculty position, A.W.L. has the following affiliations: Research Associate, National Bureau of Economic Research; Chief Investment Strategist, AlphaSimplex Group; consultant, Office of Financial Research; member, Moody’s Advisory and Academic Research Committee; member, Financial Advisory Roundtable, Federal Reserve Bank of New York; member, Economic Advisory Committee, FINRA; member, Board of Overseers, Beth Israel Deaconess Medical Center.

In addition to his MIT LFE position, R.M.S. is managing director at Moody’s Corporation.

No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No direct funding was received for this study; general research support was provided by the MIT Laboratory for Financial Engineering and its sponsors. The authors were personally salaried by their institutions during the period of writing (though no specific salary was set aside or given for the writing of this paper).

Correspondence Correspondence regarding the manuscript should be addressed to Andrew W. Lo, MIT Sloan School of Management, 100 Main Street, E62–618, Cambridge, MA 02142–1347, USA (email: alo@mit.edu). Correspondence regarding access to the megafund simulation software should be addressed to Jose-Maria Fernandez, MIT Laboratory for Financial Engineering, 100 Main Street, E62–611, Cambridge, MA 02142–1347, USA (email: jose-maria.fernandez@sloan.mit.edu)