Crises and the Direction of Innovation*

Kevin A. Bryan† Jorge Lemus‡ Guillermo Marshall§

October 8, 2020

Abstract

The value of innovation during crises can be extraordinary. While high payoffs increase the *rate* of innovation, they also induce a strategic distortion in its *direction*. High payoffs attract entry by innovators, making the R&D supply side more competitive. This competition endogenously shifts effort toward less promising but quicker-to-finish inventions. We develop a dynamic structural model quantifying the magnitude of this distortion, even when the value of potential inventions are not observed in the data. As a case study, we estimate entry of vaccines versus therapeutics and novel versus repurposed compounds developed during the Covid-19 pandemic, showing substantial distortion away from vaccines. Policy remedies include advance purchase commitments based on ex-ante value, targeted research subsidies, or antitrust exemptions for joint research ventures.

Keywords: Innovation, direction of innovation, crises, market inefficiency, pharmaceutical innovation

*We thank participants at the NBER Economics of AI conference as well as seminar participants at Universidad Adolfo Ibañez, Universidad de Chile, and University of Arizona for helpful comments. Guillermo Marshall is supported in part by funding from the Social Sciences and Humanities Research Council.

†Rotman School of Management, University of Toronto. kevin.bryan@rotman.utoronto.ca

‡University of Illinois at Urbana-Champaign, jalemus@illinois.edu

§Sauder School of Business, University of British Columbia, guillermo.marshall@sauder.ubc.ca.

1 Introduction

During a crisis, such as a war, pandemic, or environmental disaster, time is of the essence. Inventions that help solve the crisis are very valuable, especially if they are discovered quickly. Innovation policies such as patents, prizes, and research subsidies encourage entry by potential inventors. While incentivizing the entry of firms performing R&D on crisis-related inventions is important, whether these firms choose to work on socially-valuable inventions is equally important. Hence, innovation policies during a crisis should also be designed to promote an efficient allocation of research across inventions.

The intuition for why a crisis can generate an inefficient allocation of research across inventions (that is, a distortion in the direction of innovation) is as follows. In a severe crisis, the planner finds all crisis-related inventions more valuable, and needs them more quickly. To get these inventions, more firms need to perform R&D. Entry makes the supply side for research more fractured, and competition for profits may push firms in equilibrium to work on socially-inefficient solutions. Why? Each individual firm does not account for how their successful invention lowers the value of subsequent inventions by others. In particular, firms that work on low-quality inventions that can be invented quickly will not account for how they lower the ex-post value of high-quality inventions, which are partial substitutes. This negative externality implies that as a crisis grows more severe, and more firms perform R&D in response to the higher profit opportunity, firms divert effort to quicker and lower-value projects.

Our contribution is twofold. First, we present a model of invention choice formalizing the idea that a crisis can create a distortion in the direction of innovation. Second, we quantify the magnitude of this directional distortion using an empirical version of our model, which we estimate using data on project choices by pharmaceutical firms during a crisis.

The empirical challenge when estimating the distortion in the direction of innovation is that even when there are many hypothetical inventions in a given innovation race, often only a small number of them are actually invented. The value of hypothetical inventions is therefore not observed, even ex-post. Yet, the value of these uninvented inventions is critical for optimal innovation policies. To overcome this challenge, we propose a model of invention choice, which we estimate leveraging both revealed preference and data on the invention choices of firms over time. The key parameters that we estimate are the value of different inventions and relative R&D entry costs which may be heterogeneous across firms and projects. We then use the estimates of our model to quantify the magnitude of the distortion in the direction

of innovation. Further, we use our model estimates to investigate innovation policies such as advanced market commitments and R&D subsidies.

Our insight that a crisis may create directional distortions, as well as our empirical methodology to estimate this distortion applies, in principle, to study any crisis. Our empirical estimation hinges on knowing which type of firms are performing R&D, what type of R&D they are performing, when they enter, and, importantly, how severe a crisis is. To illustrate our main findings, we exploit the extensive and detailed documentation that exists for the Covid-19 crisis. We implement our empirical strategy using proprietary data on firm characteristics and project choice of entrants into Covid-19 research during the first six months of the pandemic. There are at least four reasons why the Covid-19 setting is ideal for our estimation. First, there is well-documented data on hundreds of entrants working on Covid-19 related projects. Second, there are two well-defined choices of direction such as "vaccine or therapeutic," and we know that vaccines are more difficult and more socially valuable. Third, there are standardized data on each firm's prior research capabilities. Fourth, we know exactly when the crisis started, and we observe a shock to the severity of the crisis in early March 2020 when the epidemic globalized. All these features make the Covid-19 setting ideal to explore the direction of innovation during a crisis.

To help understand the basic theoretical insight, consider the following patterns in Covid-19 related research. Figure 1 shows that the number of both pharmaceutical projects and related academic research articles were an order of magnitude more than any previous epidemic.[1] The incredible rate of invention is often seen as a success story of global drug development. As the pandemic became global in March 2020, the rate of pharmaceutical research rose yet higher.[2]

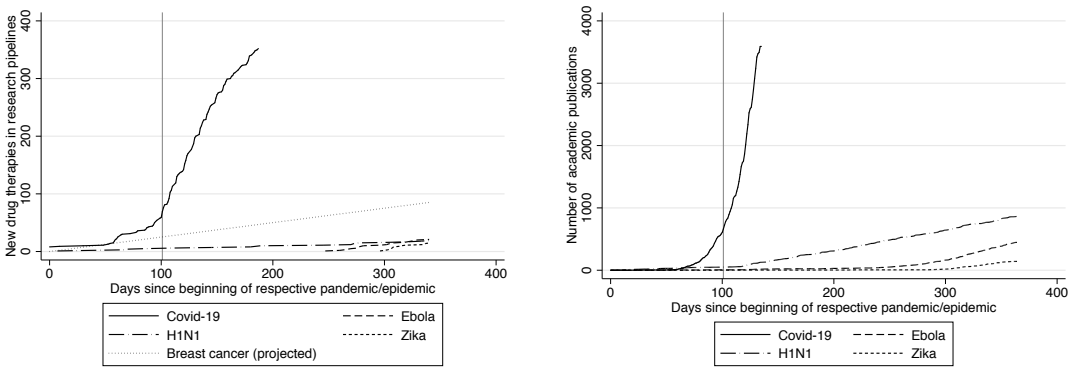
The *direction* of invention is more concerning. Compared to H1N1, Ebola, or Zika, Covid-19 research pipelines are skewed towards repurposed drugs and non-vaccine therapies.[3] Following a sharp increase in Covid-19's severity in early March 2020, Figure 2 shows that the rate of entry of vaccines did not change, even as the rate of entry on non-vaccine therapeutics soared.

[1] The 2012 MERS and 2002-2003 SARS epidemics led to only a combined 25 drug therapies in the first year after these outbreaks began. In our analysis, we use December 1st, 2019, as the start date of the Covid-19 outbreak. See, for instance, Wang et al. (2020).

[2] This is perhaps natural given the severity of Covid-19: its death toll has far exceeded recent viral outbreaks including H1N1 (2009), Ebola (2013-2016), or Zika (2015-2016). As of October 1, 2020, there have been one million confirmed deaths worldwide from Covid-19. In contrast, there were 11,323 deaths from Ebola; 18,036 deaths from H1N1, and zero reported deaths from Zika within the first year of the outbreak.

[3] Repurposed drugs are quicker to develop but are less well-targeted (Strittmatter, 2014). Vaccines are more challenging to develop than antiviral therapies (Lurie et al., 2020). Snyder et al. (2020) estimate the expected social value of an optimal vaccine development program is on the order of $3 trillion.

Figure 1: Panel A shows the number of drug therapies in pharmaceutical pipelines, by pandemic/epidemic. Panel B shows the number of disease-related academic medical publications, by pandemic/epidemic.



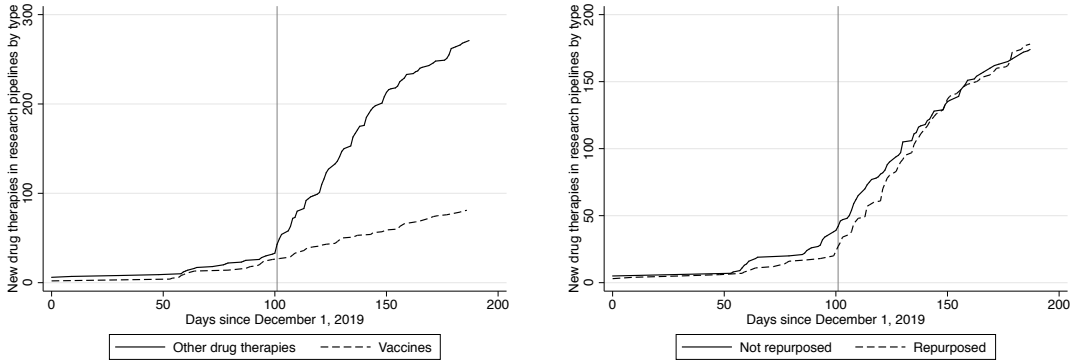
A) Therapies Pipelines B) Academic Publications

Notes: The figure plots the number of drug therapies (at all stages of development) in research pipelines, by disease. The beginning of the respective pan/epidemics are December 1, 2019 (Covid-19), April 1, 2015 (Zika), December 1, 2019 (Ebola), and January 1, 2009 (H1N1). Covid-19 therapies measured as of June 15, 2020. The projected number of breast cancer drug therapies are provided as a reference and are computed using the formula entry-rate*time, where entry-rate is the average number of new breast cancer drug therapies per day between the years 2007 and 2016. The vertical line indicates March 11, 2020, the date the WHO declared a global pandemic. See Section 3.3 for a description of the dataset.

Likewise, the post-March change in the rate of entry of novel compounds to treat Covid-19 fell relative to that of repurposed drugs. Even firms with prior experience developing vaccines became less likely to enter with a Covid-19 vaccine relative to therapeutics.

This is consistent with our theory. Firms developing therapeutics do not care that their successful invention of a decent-but-imperfect drug lowers the ex-post value of a vaccine invented later by someone else. Firms developing vaccines realize that by the time they invent, modest therapeutics are likely to have been invented by someone else. Since payoffs depend on the ex-post value of an invention, the vaccine will therefore be less valuable ex-post than ex-ante. Anecdotal evidence from top researchers taking part in an elite, international Covid recovery entrepreneurship program who were considering directing effort towards the Covid-19 research support this idea. These researchers were discouraged from entering the Covid-19 race, given the highly competitive landscape, although having top researchers working on a potential solution is arguably socially valuable (for more details, see Appendix C). Evidence from past crises also point towards the importance of the ex-post value of an invention. Indeed, "[a] decade ago, after the H1N1 influenza pandemic fizzled out, the governments of America and various European countries backed out of promised contracts, leaving pharmaceutical companies holding the bag which contained hundreds of millions of dollars of

Figure 2: Number of Covid-19 drug therapies in research pipelines, by drug classification



A) Vaccines B) Repurposed

Notes: The figure plots the number of Covid-19 drug therapies (at all stages of development) in research pipelines, by type of drug. Panel A separates the drug therapies between vaccines and all other drug therapies. Panel B separates the drug therapies between repurposed and non-repurposed drugs. Repurposed drugs are defined as drug therapies that existed prior to December 1, 2019 (i.e., beginning of the Covid-19 pandemic) and has more than one indication (e.g., Covid-19 and Ebola). See Section 3.3 for a description of the dataset.

development costs."[4] This "fizzling out" of profits is endogenous to the R&D choices made by all firms.

In the absence of directional distortions, our structural model estimates that 77 percent more firms would have worked on vaccines and 17 percent more on novel compounds between January and June 2020. A directed cost subsidy toward vaccine developers equal to 31 percent of the value of a non-vaccine drug, or 17 percent of the value of a repurposed drug for novel compound entrants, fixes the directional inefficiencies.[5] Technology-neutral advanced market commitments leave substantial directional inefficiency in both cases. Essentially, ensuring that inventors receive the full social surplus of their inventions does not solve the externality whereby firms do not account for how their successful inventions affect the value of remaining projects for other firms. Indeed, even an AMC that pays a bonus *only* for vaccines equal to their ex-post social value does not fully rectify this inefficiency: firms still react to the possibility that an adequate treatment arriving before a vaccine is invented will diminish the payoff.

Our results contribute to three long-running literatures in innovation policy. First, we contribute to the nascent literature on innovation policy and empirical direction. Budish et al.

[4] economist.com/briefing/2020/04/16/can-the-world-find-a-good-covid-19-vaccine-quickly-enough

[5] Beginning in late May 2020, large vaccine-directed subsidies began to be introduced, such as Operation Warp Speed in the United States. See https://www.hhs.gov/about/news/2020/06/16/fact-sheet-explaining-operation-warp-speed.html.

(2015) show that there is too little R&D on diseases whose necessary clinical trials are longer, because the effective term of patent protection after a drug reaches the market is therefore shorter. Moser (2005) suggests, using evidence from 19th century World's Fairs, that inventors in countries without strong patent protection shifted effort toward inventions which can be protected by secrecy, such as Swiss watches. Acemoglu (2002) develops a canonical model of directed technical change showing how factor scarcity and inventor subsidies change what is invented, while Acemoglu et al. (2012) gives an application to optimal climate change policy. Competition in the invention sector, however, is monopolistically competitive, with an implicitly fixed market structure. These studies demonstrate that R&D direction shifts to areas where inventor rewards are higher, and reacts to changes in factor prices and technological substitutability. However, to our knowledge there are no other empirical studies on distorted invention direction based on market structure, a first-order concern in crises.

Second, there is the question of how market structure relates to the *rate* of invention. Schumpeter famously argued that while competitive markets have static benefits, quasi-rents are important for covering the fixed costs of innovation.[6] The debate which builds on Schumpeter is particularly complex when market structure is affected by innovation policy. We show there is also an economically-consequential relationship between endogenous market structure and the direction of innovation. Facially-neutral policies targeted at the rate of invention, like patents, may not work as intended during crises since those policies themselves affect market structure.

Finally, our results are related to the largely historical literature about how governments consider the tradeoffs between various innovation policies during crises.[7] Prizes were common even in the 18th century particularly to solve urgent military problems like that of longitude when sailing and the canning of food for Napoleon's troops (Khan, 2015). The U.S. government formed a Patent Compensation Board in 1946 to buy out nuclear energy inventions in the early days of the Cold War (Shavell and Van Ypersele, 2001; Kremer, 1998). Patents arose to protect less-powerful inventors from Venetian politicking (Comino et al., 2020). Just as governments have used these special innovation schemes to avoid favoritism, commitment problems, and excess market power in critical markets, we argue that crises require particular attention to the effects of competitive R&D on directional incentives.

[6] See Shapiro (2011) on the "Arrow" and "Schumpeter" perspectives on this question.

[7] On the tradeoff between patents, prizes, and other inducements more generally, see, (e.g. Wright, 1983; Weyl and Tirole, 2012; Galasso, 2020).

2 A Theoretical Model of Crisis Innovation

We begin by showing the theoretical link between crises, entry incentives, and direction choice. The main insight is that intense competition among inventors induces them to "race" toward easier, lower-value inventions even when doing so is socially inefficient. Because the value of inventions ex-post depends on what substitutes exist, these low-value inventions decrease the payoff inventors working on more difficult projects with higher ex-ante value will earn. Consider the following example developing this idea in a highly restricted setting; we will loosen many simplifying assumptions in Section 2.1.

Example. There are two possible inventions. Invention A takes one year to discover and creates a social payoff of π_A. Invention B can be available immediately and generates a social payoff of π_B. After the first invention, the value of the remaining project drops to zero. Future payoffs are discounted by δ. Assume that it is socially optimal to discover A rather than B. That is,

$$\delta\pi_A > \pi_B.$$

Consider now a competitive equilibrium in which N firms can work on either A or B. If everyone works on A, each individual firm wins π_A with probability $1/N$ after one year. But if everyone works on A, a firm can deviate to work on B and immediately receive a payoff of π_B. Thus, everyone working on A is a competitive equilibrium when

$$\frac{1}{N} \cdot \delta\pi_A > \pi_B.$$

Note that with one monopolist inventor, the choice of direction is optimal. However, as competition in the research sector increases, the incentive for firms to deviate from the social optimum does as well. Firms deviate to the easier but less-valuable project because they do not account for how their invention of B means that other firms will earn nothing if they subsequently invent A.

This intuition is similar to the patent racing effect of Loury (1979). Here, instead of exerting a superoptimal rate of research effort, firms with fixed capacity shift effort to projects where the externality of their effort is larger. Directional racing is in some ways more concerning: we may not worry about superoptimal effort a la Loury if it counteracts other externalities, such as the fact that inventors capture only a portion of the social surplus of their inventions. However, firms shifting their research effort toward lower-value inventions is harder to justify

as counteracting another inefficiency.[8]

How does this externality change in a crisis? When the social value of relevant inventions in a given area is very high, governments want many firms to attempt R&D. They do so by increasing the payoffs to inventions in a given class, either directly or by subsidizing effort. Higher payoffs lead to more entry and hence stronger racing externalities. The more firms working on a problem, the greater the incentive for some of them to cut corners by working on a weaker solution which beats their rivals to market.

We will now develop this intuition formally, by expanding the model of Bryan and Lemus (2017) to a setting with endogenous entry. In particular, racing toward low-value but easy inventions occurs even when (1) the discovery time of a project is stochastic; (2) a project arrives sooner when more firms work on it; (3) successful inventions arrive in continuous time; (4) the relative arrival rates of different projects are arbitrary; (5) inventing one project does not drive the value of other projects to zero; (6) firms can allocate a portion their R&D budget to A and a portion to B (and are free to change this composition over time); and (7) the number of firms is determined endogenously, requiring fixed cost payments to enter R&D.

The model is deliberately stark, in order to clarify precisely the cause of the racing externality. Since entry depends on expected profits, and expected profits depend on post-entry equilibria, incorporating endogenous entry is not trivial.[9] When we take the model to data in the following section, we will loosen some theoretical assumptions to match important empirical features at the cost of analytic tractability. In particular, we will permit firm- and project-level heterogeneity in entry costs, and sequential entry with rational expectations.

2.1 Model

Projects: There are two projects, $j \in \{A, B\}$, characterized by three project-specific parameters: (1) the ease of invention λ_j; (2) the expected payoff $\pi_{1,j}$ to the inventing firm when nothing has been invented yet; and (3) the expected payoff $\pi_{2,j}$ to the firm that invents j if the other project has already been invented. This captures that the payoff of each project

[8] See Bryan and Lemus (2017), who identify a counteracting externality in contexts where underappropriated sequential invention is first-order. In crises, follow-on inventions may be relatively less important, so we focus our attention on directional racing.

[9] The model in this section can be extended in a number of variations while retaining analytic tractability. For example, when firm size is heterogeneous, firms with larger research capacity are less likely to deviate toward low-value, quick projects compared to smaller firms. Some extensions are available from the authors on request.

depends on the *history* of discoveries. For example, let the expected value of a vaccine A be 10 and of a treatment drug B be 5 when nothing has been invented yet. Once the vaccine is discovered, the marginal benefit of the drug falls to, say, 3, since less treatment is needed in a partially immunized population. Likewise, if the drug is invented first, the marginal benefit of the vaccine A may fall to 7 since effective therapeutics will lead only high-risk populations to vaccinate. In this case, $\pi_{1,A} = 10$, $\pi_{1,B} = 5$, $\pi_{2,A} = 7$, and $\pi_{2,B} = 3$.

Firms: Each firm is endowed with one perfectly divisible unit of effort. A firm that wants to enter the R&D race must pay a one-time fixed cost F. Each firm that enters chooses what fraction of its research capacity to allocate towards each project at each point in time. We denote by $x_{ijt} \in [0, 1]$ the research effort allocated toward project j by firm i at time t.

Timing: Time is continuous and the discount rate is $r > 0$. All firms first simultaneously choose whether to enter. Conditional on that entry, at any given time t, firms simultaneously allocate their research capacity arbitrarily across available projects.[10] The probability that firm i invents j before time t is given by an exponential distribution of parameter $\lambda_j x_{ijt}$. This implies that the research production function is constant returns to scale on a given project, both at the individual firm level and in the aggregate. The game ends after the two inventions have been discovered.

Crises: We capture the severity of the crisis by, equivalently, a scaling of all payoffs (inventions become more valuable) or a reduction in entry costs (e.g., caused by reduced regulatory burdens).[11] In more severe crises, we will show that both the planner optimal and the equilibrium number of firms entering the R&D race increases.

We make three additional assumptions to clarify precisely the nature of directional distortions. First, we assume that the payoff of any invention equals its social surplus.[12] That is, there is no gap between the surplus inventors earn and the social value of their inventions. Second, we assume that project A is difficult, yet valuable (a *long-term project*), whereas B is easier, yet less valuable (a *short-term project*). That is, $\pi_{1,A} + \pi_{2,B} > \pi_{1,B} + \pi_{2,A}$ and $\lambda_B > \lambda_A$. For instance, A may be a vaccine and B a therapeutic during an epidemic, or A

[10] Note that once A has been invented, B is the only possible project, and vice versa.

[11] For the Covid-19 crisis, there was action to cut red tape and reduce regulatory burdens: see https://www.hhs.gov/about/news/2020/03/18/hhs-takes-new-action-to-cut-red-tape-to-support-covid-19-response.html

[12] In equilibrium, scaling this payoff up or down will only affect direction choice to the extent that it changes the total number of firms that enter. To emphasize the nature of directional distortions rather than underappropriation, we therefore assume that firms fully capture the social value of their inventions. Alternatively, this can be interpreted as saying that we show the extent of directional distortion *even if* payoffs are high enough that inventors completely appropriate the expected value of their inventions.

a fully renewable power source and B a carbon emissions mitigation tool in the context of global warming.. Third, we assume that $\lambda_B \pi_{1,B} > \lambda_A \pi_{1,A}$. This implies that the *flow payoff* of the short-run project is larger than the flow payoff of the long-run project. If this were not true, then the value of the short-run project would be so low that no firm would ever work on it in equilibrium, regardless of discount rate or market structure.

2.2 Planner Optimum

Consider first the efficient allocation of research across projects by a fixed set of N firms. Following the invention of either $j \in \{A, B\}$, the planner will allocate all the research capacity towards the remaining project, denoted by $k \in \{A, B\}$, with $k \neq j$. The expected social continuation value following the invention of project j is then

$$V_j^S = \int_0^\infty \pi_{2,k} \cdot \lambda_k N e^{-\lambda_k M t} \cdot e^{-rt} dt = \frac{N\lambda_k}{r + N\lambda_k} \pi_{2,k} \tag{1}$$

In this equation, $\lambda_k N e^{-\lambda_k M t}$ is the density of the time of arrival of project k when all research capacity is allocated toward that project. Let $P_j^S = \pi_{1,j} + V_j^S$ therefore be the planner's expected payoff at the time j is invented if it is invented first. When nothing has been discovered yet, the planner chooses how to allocate effort across A and B to solve

$$\max_{(x_j)_{j\in\{A,B\}}} \int_0^\infty \sum_{j\in\{A,B\}} P_j^S \cdot \lambda_j x_j e^{-(\lambda_A x_A + \lambda_B x_B)\cdot t} \cdot e^{-rt} dt$$

subject to $x_A + x_B = N$ and $x_j \geq 0$, for $j \in \{A, B\}$. The probability that no innovation has arrived before time t is $e^{-(\lambda_A x_A + \lambda_B x_B)}$, and the rate at which project k is invented is $\lambda_k x_k$.

Lemma 1. *The planner optimally allocates all researchers to project A first and then to project B if and only if:*[13]

$$\pi_{1,A} + \pi_{2,B} \geq \pi_{1,B} + \pi_{2,A} + \frac{r}{r + N\lambda_B} \left[\pi_{2,B} - \pi_{2,A} + \left(\frac{\lambda_B - \lambda_A}{\lambda_A} \right) \pi_{1,B} \right]. \tag{2}$$

Note that for sufficiently high N, the optimal project is always the more difficult, long-term project A. Though the future is discounted, when the number of firms performing R&D is

[13] Bryan and Lemus (2017) explain the intuition for why the planner does not simultaneously research multiple projects. Intuitively, when the research production function has either constant or increasing returns to scale, there is always a "best" research line in expectation. Mathematically, the planner problem is a linear functional with linear constraints, hence the Charnes-Cooper transformation implies the optimum is a corner solution in the related linear program.

high enough, both projects can be finished arbitrarily quickly. Therefore, a planner wants firms to work on the highest value inventions ignoring their difficulty.

Let us now endogenize entry. Denote by $V(N)$ the social payoff under the efficient research direction with N firms. The optimal number of active firms, denoted by N^*, is the solution to

$$\max_{N\in\{0,1,2,\ldots\}} V(N) - F \cdot N$$

It is straightforward to show that $V(N)$ is a homogenous function in scaling all payoffs π by a constant factor. Therefore, it is equivalent to model a more severe crisis as a scaling of all payoffs π (a shock in demand that makes all inventions to become more valuable) or an equivalent reduction in entry costs F (e.g., caused by reduced regulatory burdens).

Proposition 1. *As the severity of the crisis increases,*

1. *it is optimal to increase the number of firms searching for a solution.*
2. *there exists a degree of severity such that once the crisis passes that point, it is optimal to allocate full effort toward the long-term project.*

Proposition 1 is intuitive. Crisis severity affects the direction of invention *indirectly* through the number of firms. More severe crises make it worthwhile for the planner to pay the cost of entry for more firms. As we noted above, when the number of firms who enter is sufficiently high, any project can be invented arbitrarily quickly, hence firms are optimally directed to work on high-value projects even when they are difficult.

2.3 Firm Equilibrium

In contrast to the social planner, the expected private continuation payoff after the first invention is

$$V_j = \frac{1}{N} V_j^S$$

That is, following the first successful invention, all firms can work on the remaining invention and each will invent it first with a uniform probability.

Let $a_{-ij} = \sum_{k\neq i} x_{kj}$ be the cumulative effort by firms other than i on project j, and let $P_j = \pi_{1,j} + V_j$. The best response of firm i solves

$$\max_{(x_{ij})_{j\in\{A,B\}}} \int_0^\infty \sum_{j\in\{A,B\}} (P_j \cdot \lambda_j x_{ij} + V_j \cdot \lambda_j a_{-ij}) e^{-(\lambda_A(a_{-iA}+x_{iA})+\lambda_B(a_{-iB}+x_{iB}))\cdot t} \cdot e^{-rt} dt$$

subject to $x_{iA} + x_{iB} = 1$ and $x_{ij} \geq 0$, for $j \in \{A, B\}$. The probability that no innovation has arrived before time t is $e^{-(\lambda_A(a_{-iA}+x_{iA})+\lambda_B(a_{-iB}+x_{iB}))\cdot t}$, and the rate at which project k is invented by firm i is $\lambda_k x_{ik}$. If rivals discover project k first, at rate $\lambda_k a_{-ik}$, firm i loses the immediate payoff $\pi_{1,k}$, but can still work on the remaining invention.

Lemma 2. *Suppose that the efficient research direction is project A when N firms have entered. There exists an equilibrium where all firms work on A iff*

$$\lambda_A P_A \geq \lambda_B P_B - \frac{N(\lambda_B - \lambda_A)}{r + N\lambda_A}\lambda_A P_A + N(\lambda_B \pi_{1,B} - \lambda_A \pi_{1,A}). \tag{3}$$

The condition that guarantees that A is the efficient direction is distorted by a *strategic racing externality*, captured by the term $N(\lambda_B \pi_{1,B} - \lambda_A \pi_{1,A})$. The strategic racing externality is proportional to the difference of immediate flow payoffs and strictly increasing in the number of firms. Competing firms do not internalize that, by directing their innovation effort towards the short-term project, they lower the probability that the long-term project—which is a more difficult but more socially valuable invention—is invented first by other firms. Intuitively, this is similar to business stealing from entrants with fixed costs, with the added dimension that more "business," as measured by the fraction of total surplus earned, can be stolen by deviating toward quick projects. Thus, even if the *level* of R&D is efficient, in equilibrium firms may deploy their research in an inefficient direction.

This externality has a particularly worrying consequence in very severe crises:

Proposition 2. *As the severity of the crisis increases,*

1) the number of firms that enter in symmetric equilibria weakly increases.

2) There is a cutoff crisis severity such that all firms working on the efficient, long-term project is not an equilibrium.

As a crisis becomes more severe, there is again no direct effect on direction. If the number of firms were held constant, the fact that *all* inventions in a given area see their payoffs increase by the same factor means that the optimal (or equilibrium) choice of which project to work on does not change. However, when we allow entry to be endogenous, these higher payoffs mean more firms can enter and still cover the fixed cost of performing R&D. More entry for the planner means all inventions come relatively quickly, hence it is not worth sacrificing high-value inventions for low-value but quick ones. The opposite is true in the firm equilibrium. More entry means that each firm cares more about the payoff they can get from being the

first to invent something, and less about anything invented after the first for which a given firm accrues only $\frac{1}{N}$ of the payoff in expectation.[14]

To sum things up, crises mean higher payoffs for everything invented relevant to a particular social problem. These higher payoffs mean more firms enter in equilibrium, and the planner wants lots of firms to enter. Having more firms means everything can be invented quicker, so we should work on the highest value projects. From each firm's perspective, however, more firms increases the incentive to work on quick, low-value projects by making it relatively more important to be "first" rather than "best".[15]

Finally, is there too much or too little entry overall? In general, there can be under- or over-entry in equilibrium due to two opposing forces. First, more firms means the waiting time until the first invention is shorter, hence all firms get to work on the next invention sooner: this is a positive externality, so the market solution will tend to under-supply firms. Second, when firms independently choose whether to enter, they do not account for how their entry lowers the profits captured by other firms: this is a negative externality, so the market solution will tend to over-supply firms.

Proposition 3. *In severe crises, firms will over-enter in equilibrium, relative to the optimal number of firms.*

When the entry of each firm involves business stealing, entry is more valuable to firms than to society (e.g. Mankiw and Whinston, 1986). In a crisis, the business stealing motive overwhelms the positive externality firms impose on each other by allowing each firm to work on the continuation project more quickly. Combining these propositions, with free entry and payoffs equal to the social value of any invention, in crises there will be excessive entry and the firms that enter will work on inefficiently short-term projects. Even if we manage to get the rate of entry to the optimal level, for a sufficiently severe crisis, the firms that enter have too much incentive to work on short-term, low-value projects.

[14] Part 2 in Proposition 2 holds even if the firms appropriate a fraction of the social surplus. Conditional on the number of competitors, firm choices are unaffected when payoffs are scaled down, but the equilibrium number of firms that enter decreases. A crisis that increases payoffs by a large magnitude will distort directional choices, even after accounting for fewer firms entering in equilibrium due to imperfect appropriation.

[15] Our model also allow us to study the effect of non-profit research on the direction of innovation. Many non-profit entities are trying to develop Covid-19 solutions. From the perspective of a single profit-maximizing firm, the entry of rivals exacerbates the racing effect, regardless of whether they are for profit or non-profit. Thus, the entry of large non-profit organizations in the right direction may push for-profit firms to work on short-term solutions. The Milken Institute Covid-19 treatment and vaccine tracker attempts to track not just private pharmaceutical projects, but also public studies. As of May 4, 2020, 84 percent of Covid-19 projects were wholly private or partially sponsored by the private sector.

What can be done? The theoretical solution is straightforward: increase the payoff of the long-run project relative to the short-run project, or reduce strategic racing behavior by permitting research joint ventures and similar collaborative regimes. This is standard Pigouvian economics, where we can fix an inefficiency either with taxes and subsidies, or by directly removing the externality. In the following section, we extend this model and allow for more flexibility to structurally estimate the extent of the directional inefficiency. Once we have quantitative magnitudes in hand, we will return to solutions for directional distortions in the Discussion.

3 Estimating the Magnitude of Directional Inefficiency

Consider an analyst who has access to the following data in a particular area of innovation. The analyst observes what project is chosen by each firm, the time the firm begins research on that project, and a binary variable denoting whether the firm has prior experience on similar projects of each type. The expected payoff of research of each type is unobserved, as are the fixed costs of entry which may depend on the project the firm works on and its prior experience. Even ex-post, payoffs may be unobservable both because it is their expectation rather than realization which matters for behavior, and because there are projects which may never be invented in equilibrium. We are interested in two questions. First, can we identify the planner-optimal direction choices from this data? Second, if so, and if the observed data differs from the optimum, what policies can fix it?

In order to quantitatively match real-world data, we need to extend the theoretical model in three ways. First, we will assume that firms enter sequentially rather than entering simultaneously at the beginning of the game. Second, we allow for different firm types which are defined by firm-specific prior research experience. Third, we allow the cost of working on a project to be both type- and project-specific, rather than identical for all firms and projects as in Section 2. This introduces heterogeneity in research capacity, making some firms better at research than others, but it also allows for project-specific heterogeneity (i.e., a given project may be costly for some firms even conditional on research experience).

We model project choice conditional on entry. We do not formally model the decision to enter the race, that is, firms enter at an exogenous rate in the model. While this modeling choice is a limitation, there are at least two reasons that lead us to this choice. One is that we do not have information on firms that decided not to enter the race and the other is that it allows us to remain agnostic about the mechanisms that determine the exact entry time of

a given firm. Given this assumption about entry, we model an increase in the severity of the crisis as an increase in the entry rate of firms driven implicitly by that increase in severity.[16]

Our primary quantity of interest is the fraction of firms who work on the planner-optimal project conditional on entry. To estimate this quantity, let us first describe the full empirical model, specifying which parameters are calibrated and which are free. We then discuss what variation identifies each of the free parameters. After providing descriptive evidence of the nature of entry into Covid-19 research, we describe how we will estimate the model. Finally, we will take this model to Covid-19 data, and discuss the racing behavior we estimate in that setting.

3.1 Empirical model

There is a set of potential entrants. Each entrant has a type $\theta \in \Theta$, and the distribution of types is common knowledge. Firms enter sequentially and the difference between the arrival time of two consecutive firms is $\tau_{\text{entry}} \sim \exp(\mu)$. We assume that the type and entry time of a firm are independent random variables. Upon entry, a firm chooses whether to pursue project A or project B.

The cost of pursuing project j for a firm of type θ is $c_j(\theta)$, which is a privately-observed random variable. As in Section 2, this is a one-time cost paid by the firm at the time of entering the competition. We assume that no more than $\bar{N}$ firms can enter per project.[17] Once the first invention occurs, we assume no further entry occurs, and that each firm earns a continuation value which depends on which project they began R&D on. This implicitly allows the model to capture settings where subsequent inventions of a particular type are valuable, and where follow-on inventions of the type which was not invented first retain some value. However, in our baseline model, we assume that the innovation race ends when one of the two projects is invented.

When a firm enters, the relevant state variables are the number of firms pursuing each project, (n_A, n_B). Firms are forward looking and they form beliefs about the evolution of future competition at the time of choosing what project they will work on. Note that a firm can work on only one project and this choice is irreversible. The expected value of pursuing

[16] In the context of our model, the increase in the entry rate after March 11 can be explained by a change in costs relative to the expected payoffs of inventing a drug. For instance, as argued in Footnote 11, several actions have taken place to decrease the costs of drug development for Covid-19 during the pandemic.

[17] In the Covid-19 estimation, we assume $\bar{N} = 400$, which is the 99th percentile in the distribution of number of drug projects per disease. That is, we assume that up to 800 firms can enter the innovation race.

project j conditional on the state variables (n_A, n_B) is given by

$$V^j_{n_A,n_B} = \frac{\lambda_j \pi_j + \mu\Big(E_\theta[\Pr(A|\theta, n_A, n_B)]V^j_{n_A+1,n_B} + E_\theta[\Pr(B|\theta, n_A, n_B)]V^j_{n_A,n_B+1}\Big)}{r + n_A\lambda_A + n_B\lambda_B + \mu}. \quad (4)$$

In Equation 4, firm j wins the race with flow probability λ_j, in which case it receives a payoff of π_j. With flow probability μ a new firm enters the race before a discovery has been made.[18] This new firm, depending on its type and resulting project-specific entry costs, will choose between A or B. If the new firm chooses A, the game will transition to the state (n_A+1, n_B); if the new firm chooses B, the game will transition to the state (n_A, n_B+1).

An entrant of type θ facing state variables (n_A, n_B) chooses project A when

$$V^A_{n_A+1,n_B} - c_A(\theta) > V^B_{n_A,n_B+1} - c_B(\theta). \quad (5)$$

We assume that the cumulative distribution function of the cost differences $c_A(\theta) - c_B(\theta)$ is F_θ, which is a type-specific distribution that depends on a shape parameter $\sigma(\theta)$. Thus, the entrant chooses to pursue project A with probability $\Pr(A|\theta, n_A, n_B) = F_\theta(V^A_{n_A+1,n_B} - V^B_{n_A,n_B+1})$.

Given that we impose a limit on the total number of firms that can enter each project, and that firms enter at rate μ, there is a time T such that all the firms have entered provided that no project has been invented. We can analytically compute the payoff of a firm working on project j at this time T, and use these payoffs to solve the game by backward induction.[19] We find the unique equilibrium of the game using this recursive procedure.

3.2 Identification and Counterfactuals

Assume that firms can be classified in types $\theta \in \Theta$ based on observable covariates. For example, there may be firms with observable experience in a given type of research, and those without. Let there exist one unexpected change in the severity of the crisis with a date that is known to the researcher. As discussed in Section 2, an increase in severity which increases the payoff of all inventions in a given area leads to more entry in equilibrium. We

[18] In our setting, a firm's chance of success is independent of the time the firm has been in the race, conditional on no success. Doraszelski (2003), for instance, study the impact of learning and investment on R&D races. We do not have data on investments, so we do not model this dimension.

[19] When no further entrants can enter the innovation race, the payoff of pursuing project j is given by $V^j_{\bar{N},\bar{N}} = \lambda_j \pi_j / (r + \bar{N}(\lambda_A + \lambda_B))$.

will permit the arrival rate of both types of firms to vary as a function of the severity shock.

We therefore have the following set of parameters: the project difficulties λ_A and λ_B, the discount rate r, the fixed cost variance parameters $\sigma(\theta)$ for every $\theta \in \Theta$, the payoffs π_A and π_B, the continuation values following the first invention for each firm working on a given project $\pi_{2,A}$ and $\pi_{2,B}$, the arrival rates of all firms in both periods μ_{t_1} and μ_{t_2}, and the fraction of firms of type θ in each time period $\kappa(\theta)_{t_1}$ and $\kappa(\theta)_{t_2}$.

We first normalize scale by setting $\pi_B = 1$. Hence, the payoff π_A and the cost parameters will be normalized relative to the value of the short-term invention. The discount rate is given by the modeler. Let the continuation values $\pi_{2,A}$ and $\pi_{2,B}$ be set to zero in a baseline case, or to a fixed function of π_A and π_B otherwise. For example, assume that the second invention in a given class is worth half the value of the first invention, the third is worth half that, and so on. In this case, knowing the discount rate, the arrival rates λ_j, and the estimated value π_A fully identifies $\pi_{2,A}$ and $\pi_{2,B}$.

Without observing the payoff to realized inventions, it is not possible to separately identify λ_j and π_j. Intuitively, firms may enter a given invention contest slowly because its payoff is low when invented, or because it will take a long time to invent and hence the payoff is heavily discounted. Observing the ex-post time until invention is only possible for those that are invented in equilibrium. Even there, we do not want to conflate an invention that was found quickly by good luck with the ex-ante belief by inventors that it would be so easy. We therefore choose λ_A and λ_B to match historical normal rates of development of an invention of a given type being developed by a single firm.

The remaining parameters are free. We leverage revealed preference to identify the payoff of project A, π_A, which in our empirical application is the slow-to-invent project. Although the "racing" incentive pushes firms to choose project B (the easy project), we observe firms choosing project A despite facing significant levels of competition. The one parameter in the model that can rationalize these choices is π_A. The identification of the parameters of the cost distribution of each type of firm is possible given the assumption that the value differential of choosing project A instead of B (i.e., $V^A_{n_A+1,n_B} - V^B_{n_A,n_B+1}$) does not depend on firm type. Hence, the rate at which each type of firm chooses project A, given value differentials, identifies the parameters of the cost distributions. Lastly, the identification of the parameters of the distribution of entry times or types of firms is straightforward, as these variables are readily observed in the data.

With these parameters in hand, we can examine a number of counterfactuals. First, following

the Kremer (1998) rule of thumb, we assume the social surplus of a drug is 2.7 times the fixed-price monopolist return. If firms only receive the market return on their invention, how much directional inefficiency is there? Second, if inventors receive only the market return of the short-term inventions, but the full social surplus of their long-run inventions, does that fix the inefficiency? Third, if not, how much of a relative subsidy for inventors of vaccines or novel compounds is necessary? Fourth, what level of *entry* subsidy for vaccines or novel compounds would induce enough relative entry to reach the first best? Fifth, how does the extent of directional inefficiency from Covid compare to a less severe crisis?

3.3 Descriptive Evidence from Covid-19 Pipelines

Before estimating entry into Covid-19 research structurally, we present descriptive evidence to examine the state of the Covid-19 innovation race. We use proprietary data from "BioMedTracker," a dataset produced by Informa PLC, which tracks the development history of pharmaceutical drug projects. For every pharmaceutical drug project, the dataset provides information that includes when development started, the identity of the developer, the type of drug project (e.g., vaccine or biological drug), whether it has undergone clinical trials (and when), and whether it has been approved. This information allows us to keep track of the current and past research pipelines of pharmaceutical companies. We complement these data with information from public sources, including disease-related academic publications on PubMed and information about recent viral epidemics. See Online Appendix Section A for details about the data construction.[20]

Recall from Figure 1 that the rate of Covid-19 therapies in research pipelines, and the publication rate of Covid-19 articles in academic medical journals, both exceed that of Ebola, Zika, H1N1, and even breast cancer by at least an order of magnitude.[21] Forty-one of these drug therapies were already undergoing clinical trials as of June 15, 2020 (8 are at phase I, 15 at phase II, and 18 at phase III). This exceeds the total first-year number of drug therapies for Zika and Ebola, including all those that never reached clinical trials.

Examining Figure 1 (Panel A), there is a clear visual break in the rate at which therapies entered pharmaceutical pipelines roughly 100 days after the beginning of the outbreak. This

[20] We cross-checked our data with a publicly available report by the Milken Institute on Covid-19 therapies. Both datasets track roughly the same projects in development. See the Online Appendix for more details.

[21] Cancer in general received more NIH funding than any other disease category (NIH, 2020), and breast cancer the most of any cancer type. Breast cancer is also the cancer with by far the most therapies entering clinical trials over the past quarter century (Nixon et al., 2017).

Table 1: Share of repurposed and vaccine drug therapies, by disease

	Repurposed			Vaccine			Total drug
	Count	Share	p-value	Count	Share	p-value	therapies
Covid-19	178	.506		52	.230		352
Ebola	7	.333	0.127	10	.476	.041	21
Zika	1	.066	0	12	.800	0	15
H1N1	0	.000	0	14	.737	0	19

Notes: The table displays the count and share of drug therapies that are repurposed and vaccine drug therapies in the first year after the start of the viral outbreak, by disease. p-values of two-sided tests for equality of shares (disease v. Covid-19) in 'p-value' columns.

coincides with the spread of large-scale community infection outside of Asia, the first large-scale regional lockdown outside of China (in Northern Italy, on March 8, 2020), and the global stock market decline (the Dow Jones lost nearly 1/3 of its value between March 4 and March 23, 2020). In the analysis that follows, we will delineate this increase in the severity of Covid-19 with the March 11, 2020 WHO declaration of a global pandemic.[22]

Though the rate of Covid-19 entry is very high, especially after mid-March, the type of entry shows striking patterns. Table 1 shows that Covid-19 therapies are less likely to be vaccines, and more likely to be repurposed, than therapies developed for Ebola, Zika, or H1N1.[23] Figure 2 in the Introduction shows entry by type over time.[24] The relative trend toward repurposed therapies and away from vaccines grows even stronger after the perceived severity of the pandemic increases in early March. In particular, the share of vaccines among all drug therapies is 46 percent prior to March 11, and 19 percent following the pandemic declaration. Likewise, the share of non-repurposed drugs is 64 percent prior to March 11 and 47 percent thereafter. That is, the rate of entry of vaccines is essentially constant before and after the globalization of the pandemic in mid-March, while the rate of entry of non-vaccines, especially repurposed drugs, increased dramatically.[25]

Can firm experience explain these patterns? Table 2 (Panel A) shows the comparison between the pre-existing pipeline of firms currently involved in the development of a Covid-19 drug therapy versus all other pharmaceutical firms in our dataset. Firms addressing the Covid-19

[22]Formally, a Wald supremum test identifies this structural break as occurring on March 4, 2020. Our empirical results are robust to the precise structural break date chosen.

[23]Repurposed drugs are defined as those which existed prior to the beginning of the relevant outbreak and which have multiple indications.

[24]Online Appendix Table A.2 decomposes all Covid-19 drug therapies by drug classification.

[25]Online Appendix Table A.1 shows that Covid-19 therapies are heavily concentrated among firms based in the U.S., with 60 percent based there. Despite the crisis beginning in Asia, less than 10 percent of known therapies are being led by a firm in East Asia.

Table 2: Pipeline composition of firms by involvement in other viral outbreaks

	A) Covid-19			B) H1N1			C) Ebola			D) Zika		
	Other	Entrants	Diff.	Other	Entrants	Diff.	Other	Entrants	Diff.	Other	Entrants	Diff.
Vaccine	.287	.268	-.019	.341	.438	.098	.332	.229	-.103	.317	.786	.469
			[.455]			[.028]			[.046]			[0]
Antiviral	.441	.438	-.004	.503	.738	.235	.495	.8	.305	.486	.929	.443
			[.899]			[0]			[0]			[0]
Infectious	.536	.534	-.003	.586	.777	.191	.571	.857	.286	.585	.929	.343
			[.92]			[0]			[0]			[0]
Any of above	.544	.597	.053	.59	.806	.215	.578	.863	.285	.593	.933	.34
			[.047]			[0]			[0]			[0]
Number of firms	3,549	297		1,508	58		2,234	27		2,624	22	

Notes: The table compares the pipelines of different groups of firms. The first three columns (panel A) compare the Covid-19 firms (i.e., the firms that are developing a Covid-19 drug therapy project) with all other firms (i.e., firms not developing a Covid-19 drug therapy). Other columns are defined similarly, where entrants are defined as firms that developed a drug therapy for H1N1/Ebola/Zika within a year of the start of the respective epidemic. The variables 'Vaccine', 'Antiviral', and 'Infectious' are indicators for whether a firm has had any drug therapy in its research pipeline of that type. For example, the variable Vaccine takes the value of 1 if the firm has developed vaccines in the past. 'Any of above' is an indicator that takes the value of 1 if at least one of these indicators takes the value 1. Number of firms measures the number of firms that enter into our comparison analysis. The number of firms changes across panels because in each panel we only consider the firms that had at least one drug in its pipeline one year of the start of the epidemic.

pandemic are about equally likely to have developed vaccines, antivirals, and drug therapies for infectious diseases as those not involved in Covid-19. That is, experience with similar diseases does not seem to predict entry into the race for a Covid-19 drug therapy. This was not true in previous smaller epidemics, where firms that developed a drug therapy were far more experienced both overall and in developing vaccines, antivirals, and drug therapies for infectious diseases. In addition, a comparison across panels shows that firms working on therapies for Ebola, Zika, or H1N1 were much more likely to have had experience in related diseases than the firms working on Covid-19.[26]

Just as Covid-19 entrants overall are less experienced than entrants in previous smaller epidemics, Table 3 shows that post-March 11 Covid-19 entrants have less experience with vaccines, antivirals, and with infectious diseases, and have a smaller pipeline, though they are not wholly inexperienced.[27] Firms that enter after March 11th are 17.7 percentage points more likely to repurpose therapies from their existing portfolio, and 27.8 percentage points more likely to develop non-vaccine drug therapies. That is, after the crisis became

[26] Online Appendix Table A.3 shows that this distinction holds in a probit regression of entry on experience with similar diseases even when we condition on firm size and age. For example, the table shows that prior experience developing vaccines is generally less predictive of entry into Covid-19 than entry into the other diseases.

[27] We note that the direction and statistical significance of the differences in firm observables in Table 3 are robust to dropping the five largest firms developing a Covid-19 drug therapy, which have developed more than 400 drug therapies each in the past and are currently developing a combined total of 16 drug therapies (either as the lead firm or a partner). The same holds true for the other results presented in this section.

Table 3: Entrant characteristics, by repurposed/not repurposed and entry time

	Not Repurposed	Repurposed	Diff.	Before March 11	After March 11	Diff.
Vaccine	.448	.017	-.431	.464	.186	-.278
			[0]			[0]
Repurposed	0	1	1	.357	.534	.177
			-			[.015]
Establishment year	2009.557	2005.506	-4.052	2005.571	2007.875	2.304
			[.001]			[.152]
Pipeline size	50.247	67.258	17.011	84.143	54.064	-30.079
			[.284]			[.221]
Experience w/ vaccines	.496	.096	-.401	.5	.222	-.278
			[0]			[0]
Experience w/ antivirals	.593	.32	-.272	.731	.379	-.351
			[0]			[0]
Experience w/ infectious diseases	.696	.41	-.286	.75	.49	-.26
			[0]			[0]

Notes: The table compares entrant covariates by timing of entry and whether the drug therapy is repurposed. An observation is a firm–drug therapy combination. 'Vaccine' and 'Repurposed' are indicators for whether the drug is a vaccine or a repurposed drug, respectively. 'Pipeline size' and 'Establishment year' are measures of firm size and age, respectively. The variables 'Experience w/ vaccines', 'Experience w/ infectious diseases', and 'Experience w/ antivirals' are indicators constructed based on the research pipeline of each firm. p-values of two-sided tests for equality of means in brackets.

more severe, there was more entry of small and less experienced firms, and a change in the direction of innovation towards more repurposing and non-vaccine drug therapies.

To examine whether a particular type of firm was driving the change in the direction of innovation after March 11th, we use a logistic regression to uncover the relationship between project choice and firm characteristics, controlling for whether the firm entered before or after March 11th. We estimate this regression using the full sample of Covid-19 entrants and also using the subsample of firms that have had a vaccine project and a drug project for an infectious disease prior to Covid-19 in their research pipelines (henceforth, *experienced firms*). It is reasonable to assume that these experienced firms have the know-how to develop a Covid-19 vaccine, or at least they are better equipped than firms without prior experience to handle this task.

Table 4 shows a negative and statistically significant coefficient on the post March 11 dummy, which indicates that firms that enter after March 11th are less likely to work on vaccines (all else equal). This is true for all firms (Table 4, Column 1), but also true for the subset of experienced firms (Table 4, Column 2).[28] That is, even experienced firms were shifting away from the vaccine project after March 11, which suggests that the directional change cannot be solely attributed to the increased entry rate of inexperienced firms after March 11.

[28]Online Appendix Table A.4 shows that 76 percent of experienced firms that entered before March 11 chose to develop a vaccine, while only 51 percent of experienced firms that entered after March 11 did so.

Table 4: Project choice among Covid-19 entrants: Logit regressions

	(1)	(2)
	All firms	Experienced firms subsample
	Dependent variable: Vaccine	
Post March 11	-1.071***	-1.032*
	(0.393)	(0.603)
Pipeline size	-0.000	-0.001
	(0.001)	(0.002)
Establishment year	0.074***	0.056*
	(0.023)	(0.033)
Experience w/ vaccines	3.950***	
	(0.676)	
Experience w/ infectious diseases	-0.027	
	(0.661)	
Experience w/ antivirals	-0.940	
	(0.817)	
Observations	352	80
R^2		

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. An observation is a drug project, and the outcome variable can take one of two values: vaccine or non-vaccine drug project. 'Post March 11' is an indicator that takes the value 1 if the firm's entry date is after March 11, 2020. The variable 'Pipeline size' measures the number of drug therapies that the firm has developed (active or inactive) prior to Covid-19. The variables 'Experience w/ vaccines', 'Experience w/ infectious diseases', and 'Experience w/ antivirals' are indicators constructed based on the research pipeline of each firm. The experienced firms subsample considers only firms that have had a vaccine project and a drug project for an infectious disease prior to Covid-19 in their research pipelines.

Let us sum up the descriptive evidence. The direction of Covid-19 drug development involves a much greater percentage of non-vaccine therapies and repurposed drugs than prior epidemics. This directional tilt became even stronger after the pandemic increased in severity in early March 2020. Covid-19 attracted entry from a much less specialized set of firms than prior epidemics, and again the post-March 11 entrants are yet less experienced and specialized. Intuitively, this may be seen as resulting from the higher payoffs to Covid-19 inventions attracting entry even from firms that are not experienced in antivirals, vaccines, or infectious diseases. Finally, the post-March 11 decline in vaccine and novel compound drug development is not solely the result of increased entry by less experienced firms: there is a trend away from these projects even conditional on firm experience and prior drug portfolios.

These facts are consistent with the theory in Section 2. Higher payoffs attract entry from firms that would otherwise not find the fixed cost of R&D worth paying. These firms expect

lots of competition in drug development, especially after the pandemic increases in severity. Therefore, on the margin effort shifts toward lower-value, quicker projects.

3.4 Structural Estimation

In the estimation, we assume that there are two types of firms: experienced and non-experienced, so $\Theta = \{\text{Experienced,Non-experienced}\}$. We define experienced firms as those who have had a vaccine project and a drug project for an infectious disease in their research pipelines prior to Covid-19. We set the values of the rates at which the different types of projects are invented to $\lambda_A = 0.0000555$, $\lambda_B = 0.00007607$, which are estimates based on historical data on approval times of drugs for infectious diseases, and reflect that vaccines have historically taken longer to develop (Lurie et al., 2020).[29] We set the daily discount rate to $r = 1.1^{(1/365)} - 1$, which is a common assumption in the literature. Lastly, as noted, we make a scale normalization and set the payoff of the non-vaccine drug to 1 ($\pi_B = 1$). In this way, our estimates of the payoff of a vaccine (π_A) and the costs of developing each type of project are measured relative to the payoff of a non-vaccine drug. Note that π_j are expected payoffs, so even though we impose within-class homogeneity in this expectation, the model allows different vaccines and therapeutics to vary in their ex-post value.

To capture the structural break in the entry rate of new firms, we assume that there is an exogenous (and unanticipated) change in the rate of arrival of new firms after March 11. Implicitly, this can be thought of as resulting from an unexpected shock to payoffs of all Covid-19 related inventions. Similarly, we allow for an exogenous change in the composition of potential entrants, to reflect that fewer experienced firms entered after March 11.[30] Thus, we estimate the rate of arrival of firms (μ) and the share of experienced firms (κ) both before and after March 11. Also, we assume that the cumulative distribution function of the difference in project-specific entry costs $c_A(\theta) - c_B(\theta)$ is given by $F_\theta(t) = ((t+1)/2)^{\sigma(\theta)}$ with $\sigma(\theta) > 0$ and $t \in [-1, 1]$ (i.e., cost differences cannot be greater than π_B in absolute value), and we estimate the parameters $\sigma(\theta)$ for each type.[31]

[29] Specifically, to compute λ_A and λ_B, we multiply the approval rate of drugs for infectious diseases (11.4 percent in our sample) by one over the average drug approval times of vaccines and non-vaccine drug therapies for infectious diseases. Given the empirical rate of entry into Covid-19 research, these arrival rates imply a 95 percent chance of a successful therapeutic and vaccine after 250 and 700 days of research, respectively.

[30] That is, we do not model why different types of firms enter, but instead model their project choice conditional on entering.

[31] The expected cost difference of a firm of type θ is given by $(\sigma(\theta)-1)/(1+\sigma(\theta))$. This distribution bounds the cost differences to be between -1 and 1, i.e., cost differences cannot be greater than π_B in absolute value.

In the estimation sample, a data point includes the following variables: vaccine_j (indicator for choosing project A), time to next entry_j, experienced_j (indicator for whether the firm has experience both in vaccine production and infectious diseases), Post March 11_j (indicator for whether the firm's entry time occurred after March 11), and $(n_{A,j}, n_{B,j})$ (cumulative number of entrants into projects A and B, respectively, up to that moment of time). To construct the likelihood function, we make use of Equation 5 to determine the probability that a firm of type θ facing state variables $(n_A, n_B, \emptyset)$ chooses project A as well as the parametric assumptions on the distribution of entry times (exponential distribution) and distribution of types (discrete distribution).

The probability that there is no discovery in $[0, \tau]$, the next firm enters at time τ, its type is θ, and this new entrant works on a vaccine (project A) is given by

$$e^{-(\lambda_A n_A + \lambda_B n_B)\tau} \cdot \mu e^{-\mu\tau} \cdot \kappa(\theta) \cdot F_\theta(V^A_{n_A+1,n_B} - V^B_{n_A,n_B+1}).$$

More generally, the log-likelihood function of a data point is given by

$$\begin{aligned} l_j(\delta) \quad = \quad & \text{vaccine}_j \cdot \log(F_\theta(V^A_{n_A+1,n_B} - V^B_{n_A,n_B+1})) \\ & +(1 - \text{vaccine}_j) \cdot \log(1 - F_\theta(V^A_{n_A+1,n_B} - V^B_{n_A,n_B+1})) \\ & +\text{After March } 11_j \cdot (\log(\mu_{\text{After March 11}}) - \mu_{\text{After March 11}} \cdot \text{time to next entry}_j) \\ & +(1 - \text{After March } 11_j) \cdot (\log(\mu_{\text{Before March 11}}) - \mu_{\text{Before March 11}} \cdot \text{time to next entry}_j) \\ & +\text{After March } 11_j \cdot (\text{experienced}_j \cdot \log(\kappa_{\text{After March 11}}) \\ & +(1 - \text{experienced}_j) \cdot \log(1 - \kappa_{\text{After March 11}})) \\ & +(1 - \text{After March } 11_j) \cdot (\text{experienced}_j \cdot \log(\kappa_{\text{Before March 11}}) \\ & +(1 - \text{experienced}_j) \cdot \log(1 - \kappa_{\text{Before March 11}})), \end{aligned}$$

where the value functions implicitly take into account the changes in entry rate and composition of types after March 11. The MLE estimator of the model parameters is then given by

$$\hat{\delta} = \arg\max \sum_j l_j(\delta),$$

where $\delta = (\pi_A, \mu_{\text{Before/After March 11}}, \sigma_{\text{Experienced/Non-experienced}}, \kappa_{\text{Before/After March 11}})$.

Table 5: MLE estimates of the parameters of the model

Parameter	Estimate	St. Error
π_A	40.434	15.336
π_B (normalized)	1	-
$\mu_{\text{Before March 11}}$	0.550	0.074
$\mu_{\text{After March 11}}$	3.382	0.198
$\sigma_{\text{Non-experienced}}$	3.640	0.412
$\sigma_{\text{Experienced}}$	1.004	0.208
$\kappa_{\text{Before March 11}}$	0.446	0.066
$\kappa_{\text{After March 11}}$	0.188	0.023
N	347	
$\sum_j l_j(\hat{\delta})/N$	-0.516	

Notes: Standard errors computed based on the asymptotic distribution of the MLE estimator. Calibrated parameters: $\lambda_A = .0000555$, $\lambda_B = .00007607$, and $r = 1.1^{1/365} - 1$ (time is expressed in days in the model).

4 Results

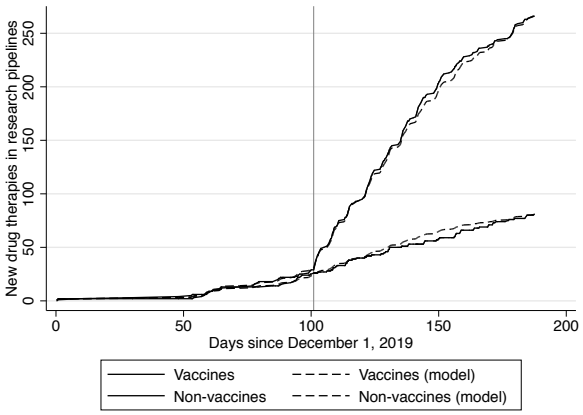
Table 5 shows the model estimates. The estimated value of a vaccine is 40.4 times the value of a non-vaccine therapy. Before March 11, the rate of entry is estimated to be 0.6 (or one firm entering on average every 1.8 days), whereas after that date, the entry rate jumped to 3.4 (or one firm entering on average every 0.3 days). Before March 11, 44.6 percent of firms were experienced, whereas the share of experienced firms dropped to 18.8 percent afterwards. The expected cost difference between the vaccine and non-vaccine projects is 0.53 and -0.07, respectively, for non-experienced and experienced firms.[32]

To gauge model fit, Figure 3 plots the raw data versus the predictions of the model for the number of vaccine and non-vaccine projects over time. We compute the model predictions by running 25,000 simulations of the model using the estimated parameters. As the figure shows, the model matches closely the number of firms in each project at every moment in time.

Using the model estimates, we also solve for the socially efficient allocation of firms across projects. In this exercise, the social planner controls whether to allocate each entrant to projects A or B, but cannot control the rate of entry of firms. When computing the planner's

[32] A motivation for estimating these parameters using revealed preference is that there are no highly-credible estimates of, for instance, the expected value of a Covid-19 therapeutic. Even when there are specifics - Gouglas et al. (2018) use confidential industry data to estimate that vaccine development between preclinical and Phase 2 trials costs an average of $31 to $68 million - mapping those estimates into our firm-specific cost estimates is not at all obvious.

Figure 3: Number of vaccine and non-vaccine drug therapies predicted by the model and in the data



Notes: Outcomes for the firm equilibrium are computed based on the average outcomes across 25,000 simulations of the game.

solution we assume that the social surplus of invention j is either equal to its private value or we use the rough estimate in Kremer (1998) that, if willingness to pay for therapies is proportional to income, the social surplus of a medical invention is 2.7x the fixed-price revenue of a monopolist inventor.[33] As before, we run 25,000 simulations of the social planner's problem and compute the outcomes for each simulation.

Table 6 presents averages across these simulations and compares them to the expected outcomes in the firm equilibrium. When assuming that the social surplus of each invention equals its private value, the analysis suggests that the social planner would increase the number of firms working on the vaccine project by 3.4 percent, or roughly three firms. However, assuming the social surplus of all Covid-19 inventions is 2.7x their private value, the planner would increase the number of firms working on the vaccine by 62 firms (77 percent).[34] Among the firms that are reallocated to work on the vaccine project, a majority of them are experienced firms.[35]

We next make use of the model estimates to quantify the impact of two sets of policy

[33]This estimate comes from assuming that willingness to pay for medical treatment is proportional to income. Using U.S. income distribution data, the gap between the total surplus of a medical invention and the profit earned by a fixed-price monopolist is 2.7.

[34]Note that we argued in Section 3 that a scaling of the payoff parameters did not affect directional choices *directly*. This was driven by the assumption that entry costs are the same for both projects. In the empirical model, however, we allow for cost heterogeneity across projects, which implies that if we scale payoffs keeping costs fixed, directional choices are affected directly.

[35]The share of experienced firms working on the vaccine project increases by 2.4 percentage points, whereas the share of non-experienced firms only increases by 1.8 percentage points.

interventions. First, we ask what directed subsidy for vaccine entrants would induce the optimal balance of vaccines and non-vaccines given estimated firm capabilities?[36] Second, we consider advanced market commitments (AMCs) which pay successful inventors the full ex-post social surplus of their inventions, or which only pay vaccine inventors that surplus while non-vaccine inventors earn only the fixed-price monopoly surplus.

Panel B of Table 6 shows that an entry cost subsidy equivalent to 1.8 percent of the value of the non-vaccine project ($\pi_B = 1$) would induce optimal direction choice if the firms are capturing the full social surplus of their inventions. Panel C shows that if the social value of Covid-19 inventions is 2.7x their private value, and hence the directional distortion is very large as noted above, a vaccine-specific entry cost subsidy equal to 31.3 percent of π_B is needed. Panel D shows that an AMC promising to pay the first inventor of any Covid-19 invention a subsidy equal to the social surplus of their invention leads to 20 percent too few firms working on vaccines.

Note why the AMC doesn't fully resolve directional distortions. Scaling up π_j, for all j, shifts some firms with high cost draws for the vaccine and low cost draws for the therapeutic to the vaccine. However, the racing externality remains: in Section 2, the derivation of directional distortions held even when firms were paid the full social surplus of their invention. Paying the AMC only if a vaccine is invented first helps, but still leaves 6.7 percent too few firms choosing to work on vaccines. To achieve efficiency with an AMC, the AMC would need to pay three times the private value of the vaccine and be paid only if a vaccine is invented first. However, achieving efficiency with an AMC is much more expensive than with directed cost subsidies ($\$121.29\pi_B$ versus $\$44.77\pi_B$). Effectively, the underprovision of vaccines is being driven by the rational expectation that some other firm will finish a moderately useful therapeutic quickly, hence large directed entry subsidies which prevent other firms from deviating are a cheaper method of preventing directional distortion.

Note where the structural model grants flexibility. Both the relative fixed costs of beginning R&D on a vaccine instead of a therapeutic, and implicitly the overall payoff of a Covid-19 related invention, are allowed to vary across firms depending on their experience with vaccines and infectious diseases. The estimates do in fact imply that payoffs were relatively high for experienced firms in the early days of the pandemic, hence their high entry share. They also imply that those firms face nowhere near the cost disadvantage of working on vaccines that

[36] Recall that firm heterogeneity is modeled by differing entry costs for vaccines for experienced firms, estimated via revealed preference. Note also that in these counterfactuals, we do not allow the number of firms who enter to vary (μ is an estimated parameter held constant in the counterfactual). The counterfactuals should therefore be interpreted as estimates for fixing directional distortion conditional on entry.

Table 6: Planner's solution versus firm equilibrium

	Number of firms working on: vaccine project (A)	non-vaccine project (B)
A. Data and model predictions		
Data	81	266
Firm equilibrium (model predictions)	80.740	266.260
B. Counterfactuals when setting the social payoff of project j to π_j		
Firm equilibrium w/ directed cost subsidy of $0.018 \cdot \pi_B$	83.472	263.528
Planner's solution	83.425	263.575
C. Counterfactuals when setting the social payoff of project j to $2.7 \cdot \pi_j$		
Firm equilibrium w/ directed cost subsidy of $0.313 \cdot \pi_B$	143.129	203.871
Planner's solution	143.048	203.952
D. Counterfactuals when making use of AMCs		
Firm equilibrium w/ AMC of $2.7 \cdot \pi_j$ for both projects	134.036	212.964
Firm equilibrium w/ AMC of $2.7 \cdot \pi_j$ for the vaccine project only	135.651	211.349
Firm equilibrium w/ AMC of $3 \cdot \pi_j$ for the vaccine project only	143.246	203.754

Notes: Outcomes are measured at 188 days since December 1, 2019. Outcomes for the firm equilibrium and planner's solution are computed based on the average outcomes across 25,000 simulations of the game. Firm equilibrium w/ directed cost subsidy indicates the case where project A receives a cost subsidy equivalent to the amount indicated in the table. Firm equilibrium w/ AMC of $2.7 \cdot \pi_j$ indicates the case when the firm inventing project j receives a payoff of $2.7 \cdot \pi_j$ instead of just π_j.

less experienced firms face.

The model does not bake in a necessary racing externality to explain the data. In Online Appendix Figure A.2, we show the results of solving the model shutting down strategic effects; that is, firms enter assuming they are the first entrant, and that no firms will enter in the future, but otherwise the model is equally flexible. The mean squared error of this model is 2.6 times higher than our full model. The difference between the full model and a nonstrategic model is especially salient when it comes to early entry. As the number of firms who have entered grows large, and no invention has yet arrived, the expected payoff for any invention becomes small due to the high level of competition. Cost differentials therefore begin to drive project choice, and hence nonstrategic models will fit well. However, when few firms have entered, an expectation that the next few firms will work on a quick therapeutic will have a substantial effect on the expected value of working on a vaccine.

Although our main estimates concern vaccines versus non-vaccines, in Online Appendix Table A.5, we replicate our analysis redefining the two possible projects to be a novel drug (project A) and a repurposed drug (project B). As in Section 2, a repurposed drug is defined as one that has more than one indication and which existed prior to the Covid-19 pandemic. A novel drug is one that is not repurposed. Based on historical data on drug approval times,

we set the values of λ_A and λ_B to 0.00006825 and 0.00009859, respectively. The value of r and the definition of experienced firms are the same as those used for the vaccine/non-vaccine drug analysis with which we lead this section. Underprovision of novel compounds is less severe than that of vaccines. A planner would increase the number of novel compounds by 1.5 percent, or by 17.0 percent if the social value of inventions were 2.7x the private value.

In Online Appendix Table A.6, we replicate our analysis reducing entry rates to 10 percent of the observed levels to measure the impact of competition on the magnitude of the directional inefficiency.[37] The table shows that there is no directional distortion when assuming that the private and social payoffs of each project are the same. When assuming that the social payoff of a project is 2.7x the private value, the planner would choose to increase the number of firms working in the vaccine project by about 5 firms (or an increase of 72 percent). Although these numbers do show that less competition decreases the directional distortion, they also show that the directional distortion is significant even with entry rates that are 10 percent of the observed ones, suggesting that our results are relevant even for innovation races that are far less crowded than the one for Covid-19.

For robustness, we replicate our main analysis assuming that the value of Covid-related inventions do not fall to zero after the first invention is found. In particular, we assume that no further entry occurs after that point, that firms can continue to work on the research line they initially entered, and that one additional invention with value $\delta\pi_j$ can be invented. That is, if $\delta = .5$, we assume that following the invention of a vaccine or a therapeutic, one additional invention still has positive value equal to half its ex-ante value. If a vaccine is invented first, this second invention can either be a second vaccine, or a first therapeutic. Likewise, if a therapeutic is invented first, this second invention can either be a second therapeutic or a first vaccine. Online Appendix Table A.7 shows that permitting a further invention causes only a small change in the quantitative welfare cost of directional distortion. Note that this setting is identical to a setting where an infinite number of inventions of either type have value, with the value of each consecutive invention being worth a fraction $\delta\frac{r}{r+n_A\lambda_A+n_B\lambda_B}$ of the prior one.

[37] Lower entry rates would be explained by an increase of entry costs relative to the expected payoffs of the different projects. Although we do not formally model the decision to enter the race (i.e., we only model project choice conditional on entry), in these exercises we scale down the estimated values of π_A and π_B to capture the reduced entry incentives that would lead to a drop in entry to 10 percent of the observed level. See the footnote of Online Appendix Table A.6 for details.

5 Discussion and Implications

We show that crises, by increasing the payoff of relevant inventions, distort the direction of research by endogenously affecting market structure. Higher payoffs cause more firms to begin research on a particular problem. The more fractured the market for research, the less any one firm weighs the total value of the market against the profits earned by racing to enter the market first with a mediocre solution. Therefore, strategic interaction in the highly competitive crisis market for invention leads to too much work on "quick" projects like repurposed drugs and too little work on long-run projects like vaccines.

We present a structural model to estimate the size of this directional distortion. Although our economic insight and structural model can be applied, in principle, to any crisis, we take advantage of the wealth of documentation that exists for the Covid-19 crisis. The richness of the data make the Covid-19 an ideal setting to apply our model and to empirically estimate the distortion in the direction of innovation. In our estimation, we use proprietary data on research pipelines during the first six months of the Covid-19 pandemic. Although the rate of Covid-19 research is proceeding at an unprecedented historical pace, it has involved more research on short-term solutions than previous epidemics. This trend was exacerbated when Covid-19 became a global pandemic in March 2020.

The distortion we identify is *ex-ante*: it shifts what firms invent rather than the deadweight loss they generate ex-post if, for example, they have a patent. This distortion holds even when we assumed that the payoff to inventors was exactly equal to the social surplus of their invention. Prior research suggests that vaccines may be underprovided due to the difficulty of extracting this surplus. For instance, Kremer and Snyder (2015) argue that for rare but serious diseases, when consumer valuation for treatment is highly heterogeneous, it is easier to extract this surplus with treatments than vaccines. We have assumed away this possibility by giving inventors the entire social surplus, yet vaccines are still underprovided. Vaccines may also be underprovided due to a commitment problem on the part of the government. Kremer et al. (2020) argue for advance purchase commitments partly on these grounds: pharmaceutical firms otherwise believe their vaccines will be expropriated after the cost of research has been incurred.[38]

These are important concerns. Nonetheless, we suggest that policymakers during a crisis concern themselves not only with the size of the payoff ex-post inventors receive, but also

[38] In an early book on the economics of Covid-19, Gans (2020) discusses in more depth the use of AMCs and the problem of commitment in previous epidemics.

on how the land rush into crisis R&D affects which projects inventors will pursue. A firm that produced a foolproof Covid-19 vaccine or a partially effective therapeutic at the peak of the pandemic would have countries bidding richly for the first batches. Nonetheless, if so many firms are working on the therapeutic that it surely will appear within a few months, potential vaccine inventors may rationally abandon that long-run project under the correct belief that the vaccine will have lower ex-post value. Technologically neutral policy in the face of strategic behavior is not in fact neutral.

Our model, both theoretical and structural, applies to problems beyond Covid-19. Consider a wartime government hoping to incentivize new aircraft, or an IGO that wants to see effective novel climate change mitigation technology. Assume that they credibly commit to pay the full, ex-post social value of any completed invention within their bailiwick. Doing so may be *worse* than paying lower rewards if the induced competition pushes firms to work on second-best technology that can be completed quickly. Our structural model permits retrospective analyses of these distortions even when the value of different inventions is unknown to the analyst, since we draw on the revealed preference of inventors to infer the magnitude of any distortions.

What can be done? Patent buyouts (Kremer, 1998), where the government buys a patent in order to remove the deadweight loss of monopoly pricing, do not solve our problem. Indeed, by increasing the return of crisis inventions, it induces more entry and makes directional distortion worse. The same is true of generic research subsidies. The fundamental problem is that the government needs to simultaneously induce entry *and* prevent the firms that enter from deviating to quick, low-value projects.

Three policies will limit this problem. First, the government can allow research joint ventures without antitrust restrictions. Note from the model that larger firms are less likely to deviate, and that the total social surplus for the industry is decreasing in the extent of deviation. In April 2020, Sanofi and GlaxoSmithKline, normally rivals, formed a joint research venture to develop a Covid-19 vaccine.[39] Research joint ventures on projects that are expected to be harder to invent that most inventions in a sector ought to be encouraged.[40]

Second, targeted subsidies, incentivizing only difficult, high-value remedies to the crisis, while permitting unsubsidized research on other projects, simultaneously induce entry and incentivizes firms to work on socially optimal projects. Targeted subsidies generally have a bad

[39] See https://www.gsk.com/en-gb/media/press-releases/sanofi-and-gsk-to-join-forces-in-unprecedented-vaccine-collaboration-to-fight-covid-19/ for details.

[40] See Grossman and Shapiro (1986) for a deeper analysis of the antitrust issues with research collaborations.

name: many innovation scholars do not like the government to "pick winners". In crises, however, the nature of high-value inventions is often widely known. There is no ambiguity, for instance, about the therapeutic properties of the highest-value remedies for Covid-19. To this end, Table 6 shows that a vaccine-specific cost subsidy of 31 percent of the value of non-vaccines ($\pi_B = 1$) would fix the directional inefficiency. We use data only through June 15, 2020 because targeted subsidies toward vaccines became a major part of the policymaker arsenal with the announcement of "Operation Warp Speed" subsidies. This shift was not ex-ante obvious: indeed, a group of prominent economists argued in the May 4th, 2020 issue of the New York Times discussed existing large subsidy programs for Covid-19 inventions, but noted that they were targeted broadly at "diagnostics, therapeutics, and treatments" (Athey et al., 2020).

Finally, advanced market commitments (AMCs) can be used, with a twist. The reason firms deviate to short-run solutions is partly because the *marginal* value of the ex-ante best project falls once partial remedies exist. This collapse may not be linear. For instance, imagine that a partially effective drug is half as good as a vaccine from the perspective of a government. Once the drug exists, firms will consider whether to keep working on the vaccine and receiving this lower payment, or to work on some outside option. Reasoning that when other firms are working on the drug, the probability of finishing a vaccine first is quite low, all firms will eventually deviate to working on the drug, and the vaccine will not be invented. An AMC committing to pay the *ex-ante* social value of an invention, even if future inventions lower their value, can completely remedy directional inefficiency. Estimates of the value of targeted Covid-19 vaccine AMCs argue that an advance commitment of nearly $40 billion, with coordinated allocation to high risk populations, increases welfare by avoiding ex-post bidding wars for potentially limited vaccine supplies (Snyder et al., 2020). Our results suggest that optimal vaccine AMCs, or alternatively vaccine cost subsidies, need to be higher yet given the strategic distortions induced by high overall Covid-19 payoffs.

6 References

Acemoglu, Daron (2002) "Directed Technical Change," *Review of Economic Studies*, Vol. 69.

Acemoglu, Daron, Phillipe Aghion, Leonardo Bursztyn, and David Hemous (2012) "The Environment and Directed Technical Change," *American Economic Review*, Vol. 102.

Athey, Susan, Michael Kremer, Christopher Snyder, and Alex Tabarrok (2020) "In the Race for a Coronavirus Vaccine, We Must Go Big. Really, Really Big," *The New York Times*, May 4 edition.

Bryan, Kevin A and Jorge Lemus (2017) "The direction of innovation," *Journal of Economic Theory*, Vol. 172, pp. 247–272.

Budish, Eric, Benjamin N Roin, and Heidi Williams (2015) "Do firms underinvest in long-term research? Evidence from cancer clinical trials," *American Economic Review*, Vol. 105, pp. 2044–85.

Comino, Stefano, Alberto Galasso, and Clara Graziano (2020) "Market Power and Patent Strategies: Evidence from Renaissance Venice," *Journal of Industrial Economics*, Vol. forthcoming.

Doraszelski, Ulrich (2003) "An R&D race with knowledge accumulation," *Rand Journal of economics*, pp. 20–42.

Galasso, Alberto (2020) "Rewards versus intellectual property rights when commitment is limited," *Journal of Economic Behavior & Organization*, Vol. 169, pp. 397–411.

Gans, Joshua (2020) *Economics in the Age of COVID-19*: MIT Press.

Gouglas, Dimitrios, Tung Thanh Le, Klara Henderson, Aristidis Kaloudis, Trygve Danielsen, Nicholas Caspersen Hammersland, James M Robinson, Penny M Heaton, and John-Arne Røttingen (2018) "Estimating the cost of vaccine development against epidemic infectious diseases: a cost minimisation study," *The Lancet Global Health*, Vol. 6, pp. e1386–e1396.

Grossman, Gene M and Carl Shapiro (1986) "Research Joint Ventures: An Antitrust Analysis," *Journal of Law, Econ. & Org.*, Vol. 2, p. 315.

Khan, B Zorina (2015) "Inventing prizes: a historical perspective on innovation awards and technology policy," *Business History Review*, Vol. 89, pp. 631–660.

Kremer, Michael (1998) "Patent buyouts: A mechanism for encouraging innovation," *The Quarterly Journal of Economics*, Vol. 113, pp. 1137–1167.

Kremer, Michael, Jonathan D Levin, and Christopher M Snyder (2020) "Advance Market Commitments: Insights from Theory and Experience,"Technical report, National Bureau of Economic Research.

Kremer, Michael and Christopher M Snyder (2015) "Preventives versus treatments," *The Quarterly Journal of Economics*, Vol. 130, pp. 1167–1239.

Loury, Glenn C (1979) "Market Structure and Innovation," *The Quarterly Journal of Economics*, pp. 395–410.

Lurie, Nicole, Melanie Saville, Richard Hatchett, and Jane Halton (2020) "Developing COVID-19 vaccines at pandemic speed," *New England Journal of Medicine*.

Mankiw, N Gregory and Michael D Whinston (1986) "Free entry and social inefficiency," *The RAND Journal of Economics*, pp. 48–58.

Moser, Petra (2005) "How do patent laws influence innovation? Evidence from nineteenth-century world's fairs," *American economic review*, Vol. 95, pp. 1214–1236.

NIH (2020) "Estimates of Funding for Various Research, Condition, and Disease Categories," https://report.nih.gov/categorical_spending.aspx, Accessed: 2020-04-23.

Nixon, Nancy A, Omar F Khan, Hasiba Imam, Patricia A Tang, Jose Monzon, Haocheng Li, Gavin Sun, Doreen Ezeife, Sunil Parimi, Scot Dowden et al. (2017) "Drug development for breast, colorectal, and non–small cell lung cancers from 1979 to 2014," *Cancer*, Vol. 123, pp. 4672–4679.

Shapiro, Carl (2011) "Competition and innovation: did arrow hit the Bull's eye?" in *The Rate and Direction of Inventive Activity Revisited*: University of Chicago Press, pp. 361–404.

Shavell, Steven and Tanguy Van Ypersele (2001) "Rewards versus intellectual property rights," *The Journal of Law and Economics*, Vol. 44, pp. 525–547.

Snyder, Christopher M, Kendall Hoyt, Dimitrios Gouglas, Thomas Johnston, and James Robinson (2020) "Designing Pull Funding For A COVID-19 Vaccine: Article proposes a âĂIJpullâĂİ program that incentivizes late-stage development and manufacturing of COVID-19 vaccines by awarding advance purchase commitments to bidding firms.," *Health Affairs*, pp. 10–1377.

Strittmatter, Stephen M (2014) "Overcoming drug development bottlenecks with repurposing: old drugs learn new tricks," *Nature medicine*, Vol. 20, pp. 590–591.

Wang, Chen, Peter W Horby, Frederick G Hayden, and George F Gao (2020) "A novel coronavirus outbreak of global health concern," *The Lancet*, Vol. 395, pp. 470–473.

Weyl, E Glen and Jean Tirole (2012) "Market power screens willingness-to-pay," *The Quarterly Journal of Economics*, Vol. 127, pp. 1971–2003.

Wright, Brian D (1983) "The economics of invention incentives: Patents, prizes, and research contracts," *The American Economic Review*, Vol. 73, pp. 691–707.

Online Appendix

Crises and the Direction of Innovation

Supplemental Material – Intended for Online Publication

A Data Sources and Data Construction

We use proprietary data from "BioMedTracker," which is an Informa PLC product and tracks pharmaceutical pipelines over time. We also retrieved lists of medical research articles by disease from PubMed to study the evolution of the volume disease-related academic publications around the time of an epidemic/pandemic.

We use BioMedTracker (last accessed June 15, 2020) to obtain the full list of Covid-19 drug therapies in development as well as the development history (i.e., the start dates of development and clinical trials if applicable) and the list of companies involved in the development of each drug therapy. Similarly, we use BioMedTracker to obtain the same information for the H1N1 pandemic (2009), the Ebola epidemic (2013-2016), and the Zika epidemic (2015-2016). We also use BioMedTracker to obtain the pipelines (i.e., the list of all drugs that are currently in development or have been in development in the past) of all pharmaceutical companies.

With few exceptions, the variables we use in the analysis are variables that are available in the raw BioMedTracker data. We define the variable 'Repurposed,' as any drug for disease x that existed prior to the beginning of the epidemic of disease x (e.g., a repurposed Covid-19 drug is one that has multiple indications and existed prior to December 1, 2019). We also define variables related to the drug-development experience of firms (i.e., "experience w/ vaccines", "experience w/ antivirals", and "experience w/ infectious diseases"), which are based on the research pipeline of each firm.

There are, of course, many other datasets on Covid-19 projects. Hand-checking these data reveal that they generally overlap heavily with the BioMedTracker data. For instance, the Milken Institute Covid tracker based on public media reports as of April 20, 2020, finds 146 drug treatments and 92 candidate vaccines, of which 49 are not modified existing platforms.[41] As of April 20, 2020, BioMedTracker finds 170 drug treatments and 51 candidate vaccines. For reasons of comeasurability with the Ebola, Zika, and H1N1 data, we use only the remedies in the BioMedTracker dataset.

[41] See https://milkeninstitute.org/covid-19-tracker.

B Additional Tables and Figures

Table A.1: Covid-19 firms, by country

Country	Freq.	Percent	Cum.
Australia	4	1.14	1.14
Austria	4	1.14	2.28
Belgium	4	1.14	3.42
Canada	19	5.41	8.83
China	10	2.85	11.68
Denmark	1	0.28	11.97
France	10	2.85	14.81
Germany	7	1.99	16.81
India	2	0.57	17.38
Ireland	3	0.85	18.23
Israel	8	2.28	20.51
Italy	4	1.14	21.65
Japan	10	2.85	24.50
Korea (South)	10	2.85	27.35
Netherlands	4	1.14	28.49
Norway	1	0.28	28.77
Russia	2	0.57	29.34
Scotland	1	0.28	29.63
Spain	5	1.42	31.05
Sweden	5	1.42	32.48
Switzerland	12	3.42	35.90
Taiwan	1	0.28	36.18
Turkey	1	0.28	36.47
United Kingdom	12	3.42	39.89
United States	211	60.11	100.00

Notes: The table shows the distribution of locations of Covid-19 firms (i.e., the firms that are leading a Covid-19 drug therapy project).

Table A.2: Covid-19 drug therapies, by drug classification

Type	Not repurposed	Repurposed	Total
Biologic	60	62	122
Device	0	1	1
New Molecular Entity (NME)	19	99	118
Non-NME	12	13	25
Unknown	2	0	2
Vaccine	78	3	81
Total	171	178	349

Notes: The table shows the number of new Covid-19 drug therapies (at all stages of development) by drug classification for both repurposed and non-repurposed drugs. Repurposed drugs are defined as drug therapies that existed prior to December 1, 2019 (i.e., beginning of the Covid-19 pandemic) and has more than one indication (e.g., Covid-19 and Ebola).

Table A.3: Probability of entry on pipeline composition: Probit regressions

	Covid-19	Ebola	H1N1	Zika
Experience w/ vaccines	0.314***	0.346*	0.824***	1.255***
	(0.120)	(0.208)	(0.212)	(0.223)
Experience w/ antivirals	0.555***	0.500*	0.630**	4.419***
	(0.134)	(0.260)	(0.268)	(0.163)
Experience w/ infectious diseases	0.108	0.632**	0.087	-3.595***
	(0.125)	(0.298)	(0.243)	(0.251)
Controls:		Firm age, firm size		
Observations	3475	2237	1516	2625
R^2				

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The table shows estimates predicting entry into the respective diseases, considering the full set of firms that have been active up until two years after the start of the respective viral outbreak. The variables 'Experience w/ vaccines', 'Experience w/ antivirals', and 'Experience w/ infectious diseases' are indicators for whether a firm has had any drug therapy in its research pipeline of that type prior to the respective epidemic. For example, the variable 'Experience w/ vaccines' takes the value of 1 if the firm has developed vaccines in the past. Firm size is defined as the number of drug therapies that the firm has developed (active or inactive).

Table A.4: Project choice among Covid-19 entrants (experienced firms subsample)

	Before March 11	After March 11	Total
Non-vaccine	6	27	33
Vaccine	19	28	47
Total	25	55	80

Notes: An observation is a drug project, and the outcome variable can take one of two values: vaccine or non-vaccine drug project. Experienced firms are the firms that have had a vaccine project and a drug project for an infectious disease prior to Covid-19 in their research pipelines. 'Before/After March 11' are indicators that take the value 1 if the firm's entry date is after March 11, 2020.

Table A.5: Planner's solution versus firm equilibrium: Repurposed vs. non-repurposed drugs

	Number of firms working on:	
	non-repurposed project (A)	repurposed project (B)
A. Data and model predictions		
Data	169	178
Firm equilibrium (model predictions)	169.562	177.438
B. Counterfactuals when setting the social payoff of project j to π_j		
Firm equilibrium w/ directed cost subsidy of $0.015 \cdot \pi_B$	172.139	174.861
Planner's solution	172.095	174.905
C. Counterfactuals when setting the social payoff of project j to $2.7 \cdot \pi_j$		
Firm equilibrium w/ directed cost subsidy of $0.17 \cdot \pi_B$	198.873	148.127
Planner's solution	198.428	148.572

Notes: The estimates of the parameters of the model are $\hat{\pi}_A = 17.857$, $\sigma_{\text{Non-experienced}} = 1.443$, $\sigma_{\text{Experienced}} = 0.389$, and the parameter estimates of the entry rate of firms and the distribution of firm types are identical to those in Table 5. The values of λ_A, λ_B, and r are set at 0.00006825, 0.00009859, and $1.1^{1/365} - 1$, respectively. As in Table 5, the values of λ_j are calibrated based on historical data on drug approval times. The definition of experienced firms are identical to those used in the vaccine/non-vaccine drug analysis in Table 5. Outcomes are measured at 188 days since December 1, 2019. Outcomes for the firm equilibrium and planner's solution are computed based on the average outcomes across 25,000 simulations of the game.

Table A.6: Planner's solution versus firm equilibrium when reducing firm entry by 90 percent

	Number of firms working on:	
	vaccine project (A)	non-vaccine project (B)
Firm equilibrium (model predictions)	7.065	25.935
Planner's solution (social payoff$_j = \pi_j$)	7.211	25.789
Planner's solution (social payoff$_j = 2.7 \cdot \pi_j$)	12.136	20.864

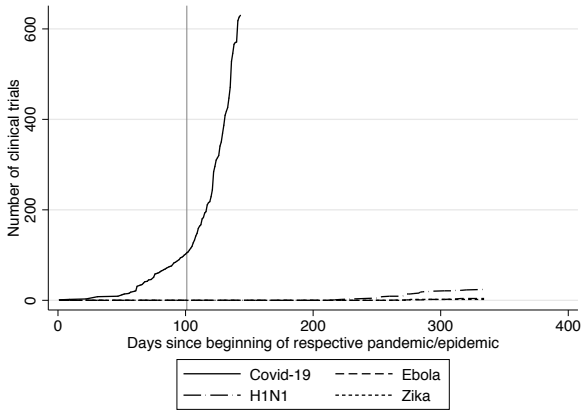
Notes: Simulations are based on the estimates in Table 5. Entry rates both before and after March 11 are set at 10 percent of the estimated values in Table 5. We also scale down π_A and π_B (relative to the entry costs) to make entry less attractive to the point that only 10 percent of firms want to enter the race. To do this, we compute an upper bound on the entry cost of each firm using the condition $V_{ij} > c_{ij}$, where j is the project chosen by firm i and V_{ij} is the value of firm i when pursuing j in the observed equilibrium (given state variables). Assuming that all firms that chose project j faced an entry cost given by $\min_i c_{ij}$, we search for the lowest scalar x such that $V_{ij} > x * c_{ij}$ holds for only 10 percent of firms. We then use x to scale down π_A and π_B in these counterfactuals. All other estimated parameters remain unchanged. Outcomes for the firm equilibrium and planner's solution are computed based on the average outcomes across 25,000 simulations of the game.

Table A.7: Planner's solution versus firm equilibrium when allowing for two consecutive races

	Number of firms working on:	
	vaccine project (A)	non-vaccine project (B)
A. $\delta = 0.1$		
Data	81	266
Firm equilibrium (model predictions)	80.893	266.107
Planner's solution (social payoff$_j = \pi_j$)	83.608	263.392
Planner's solution (social payoff$_j = 2.7 \cdot \pi_j$)	144.513	202.487
B. $\delta = 0.5$		
Data	81	266
Firm equilibrium (model predictions)	80.915	266.085
Planner's solution (social payoff$_j = \pi_j$)	83.706	263.294
Planner's solution (social payoff$_j = 2.7 \cdot \pi_j$)	130.693	216.307

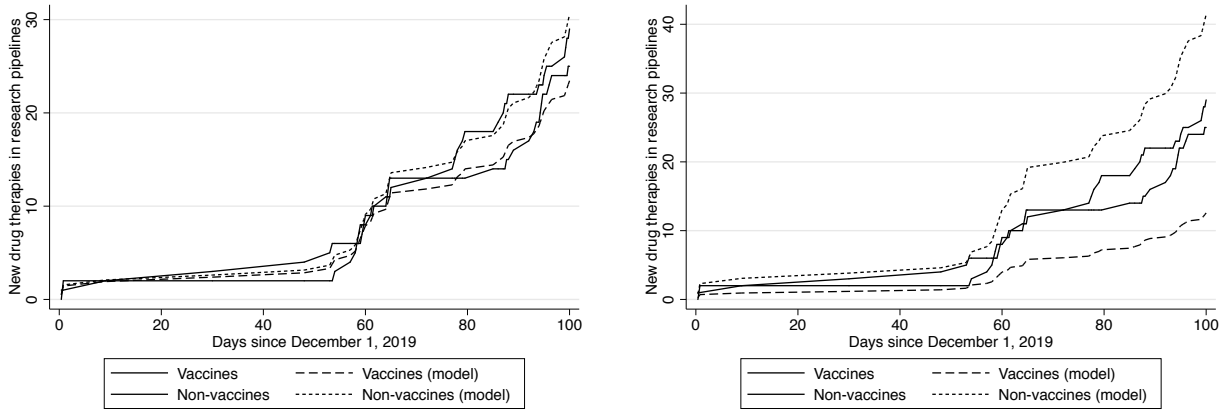
Notes: The estimates of the parameters of the model in Panel A are $\hat{\pi}_A = 37.126$, $\sigma_{\text{Non-experienced}} = 3.643$, and $\sigma_{\text{Experienced}} = 1.005$; the in Panel B are $\hat{\pi}_A = 28.072$, $\sigma_{\text{Non-experienced}} = 3.650$, and $\sigma_{\text{Experienced}} = 1.007$. The parameter estimates of the entry rate of firms, the distribution of firm types, and the calibrated parameters are identical to those in Table 5. Outcomes are measured at 188 days since December 1, 2019. Outcomes for the firm equilibrium and planner's solution are computed based on the average outcomes across 25,000 simulations of the game.

Figure A.1: Number of clinical trials registered with clinicaltrials.gov, by pandemic/epidemic



Notes: The figure plots the number of clinical trials (all stages) during the first year after the start of the viral outbreak, by disease. We use data from clinicaltrials.gov. We note that a given drug therapy may undergo multiple clinical trials. The beginning of the respective pan/epidemics are December 1, 2019 (COVID-19), April 1, 2015 (Zika), December 1, 2019 (Ebola), and January 1, 2009 (H1N1). COVID-19 therapies measured as of April 22, 2020. The vertical line indicates March 11, 2020.

Figure A.2: Number of vaccine and non-vaccine drug therapies predicted by the model and in the data



A) Strategic model B) Myopic model

Notes: Outcomes for the firm equilibrium are computed based on the average outcomes across 25,000 simulations of the game. The strategic model in panel A corresponds to the model presented in Section 4. The myopic model in panel B corresponds to a version of the model in Section 4 in which each firm behaves as if it is the only firm that has entered and will ever enter the race. The figures restrict attention to the first 100 days of the pandemic.

C Competition and Directional Choice: Anecdotal Evidence from Expert Interviews

To what extent do we observe this racing behavior directly? Firms generally do not make their rationale for choosing one R&D project over another observable to the analyst. However, we do see suggestive evidence of our mechanism at play in qualitative data from a four-month Covid-related entrepreneurship program run in the Spring and Summer of 2020, which for anonymity reasons we call the Program. In the Program, 65 science-based startups from around the world participated in monthly, structured online meetings with a panel of entrepreneurship experts which included serial founders, partners at leading venture capital firms, and world-renowned scientists and epidemiologists. Many of the founders had deep technical expertise but little business experience. At each meeting, the panel gave firms advice on the long-run financial viability of their Covid-related business. We therefore can observe, qualitatively, the interaction between technical potential and financial viability. Although these companies are largely not pharmaceutical companies, the tradeoff of being "first" versus being "best" came up frequently in the online meetings. Consider the following cases.

One firm, headed by senior academics at a top global research university, had developed a technique for a new type of vaccine which is particularly promising when it comes to coronaviruses in particular. In an early summer meeting, the former director of a large government health body evaluated the firm as a "brilliant company. Currently, lots of competitors in vaccine space, but this approach is so far superior." By August, a half dozen different mentors told the company the vaccine space was simply too competitive for them to succeed even if their approach was superior: "This is a very competitive space. Point of differentiation in relation to competitors is very important. There is a race to the forefront in this crowded space." The advice in the final meeting was to either stop working on Covid-19 altogether and focus on a broader scientific problem, or to license out any aspects of the Covid-19 technology which can speed up development for a more advanced competitor.

A second firm, also founded by academics at a top global research university, produced a sensor which could identify pathogens on surfaces or in water. Their technology could detect contaminated surfaces on-site, without the use of specialized equipment or trained clinicians. Although the mentors "like the team" and believe "there is good technical knowledge," even by June there was widespread agreement that "this is becoming a crowded space, so understanding the current players/emerging competition is critical." Even though the firm's technology was quicker and more accurate than existing competitors, its cost and time to

commercial development were too high compared to easier-to-develop technologies such as strong surface cleaners. The company pivoted away from Covid-19 development. Note that in both cases, there was general agreement that the proposed invention was the leader on technical grounds. However, there was so much competition in both markets that modest partial substitutes which arrived to market first were expected to take much of the profit.

D Analysis of the Model

Optimal Direction (Lemma 3).

Lemma 3. *For any fixed capacity N:*

1. *The planner optimally uses all the research capacity on project A first and then on project B if and only if $S_A(N) \geq S_B(N)$ where*

$$S_j(N) = \frac{N\lambda_j}{r + N\lambda_j} P_j. \tag{6}$$

2. *The condition $S_A(N) \geq S_B(N)$ is equivalent to*

$$\lambda_A P_A \geq \lambda_B P_B - \Delta(N)\lambda_A P_A, \tag{7}$$

where $\Delta(N) = \frac{N(\lambda_B - \lambda_A)}{r + N\lambda_A}$.

3. *The condition $S_A(N) \geq S_B(N)$ is also equivalent to*

$$\pi_{1,A} + \pi_{2,B} \geq \pi_{1,B} + \pi_{2,A} + \underbrace{\frac{r}{r + N\lambda_B}\left[\pi_{2,B} - \pi_{2,A} + \left(\frac{\lambda_B - \lambda_A}{\lambda_A}\right)\pi_{1,B}\right]}_{\equiv g(N)}, \tag{8}$$

where $g(N) \geq 0$, for all N, and $g(\cdot)$ is strictly decreasing with $g(N) \to 0$ as $N \to \infty$.

Proof: 1. and 2. Directly from Bryan and Lemus (2017).

3. The condition $S_A \geq S_B$ can be written as

$$\pi_{1,A} + \pi_{2,B} \geq \pi_{1,B} + \pi_{2,A} + g(N)$$

where $g(N)$ is strictly decreasing and converges to zero as $N \to \infty$, under our assumptions. Note that when $N \to \infty$, $\frac{N\lambda_B}{r + N\lambda_B} \to 1$, so the condition above becomes $\pi_{1,A} + \pi_{2,B} \geq \pi_{1,B} + \pi_{2,A}$.

Efficient Entry in a Crisis (Proposition 1).

1. Consider a crisis of severity β, where β multiplicatively scales all payoffs π. Then, the optimal number of firms is

$$\max_{N \in \{0,1,\ldots\}} V(N)\beta - F \cdot N$$

Given that $V(N)$ is an increasing function (it is the maximum of two increasing functions), a direct application of Topkis' Theorem implies that N^* is weakly increasing in β.

2. As $\beta \to \infty$, we have that $N^*(\beta) \to \infty$. This implies that $\frac{N^*(\beta)\lambda_B}{r+N^*(\beta)\lambda_B} \to 1$. Simple algebra shows that the condition becomes $\pi_{1,A} + \pi_{2,B} \geq \pi_{1,B} + \pi_{2,A}$.

Equilibrium (Lemma 2).

This result is a direct from Corollary 1 in Bryan and Lemus (2017).

Entry and Direction in Equilibrium (Proposition 2)

Part i: Let β scale multiplicatively all payoffs. Let Π^e represent equilibrium profits in a symmetric equilibrium. The equilibrium number of firms N^e in a crisis of severity β is determined by the condition

$$\Pi^e(N^e) \geq \frac{F}{\beta} > \Pi^e(N^e + 1).$$

Given that $\Pi^e(\cdot)$ is weakly decreasing, the equilibrium number of firms increases with the severity of the crisis.

Part ii: Note that as $N^e \to \infty$ we have $P_A \to \pi_{1,A} + \pi_{2,B}$, $P_B \to \pi_{1,B} + \pi_{2,A}$, $\Delta(N) \to \frac{\lambda_B - \lambda_A}{\lambda_A}$. Therefore, for N^e large enough we will have

$$\lambda_A P_A < \lambda_B P_B - \Delta(N)\lambda_A P_A + N(\lambda_B \pi_{1,B} - \lambda_A \pi_{1,A}).$$

Excessive Entry (Proposition 3)

The marginal condition that determine the efficient number of firms to enter (ignoring the integer constraint) is $G(N) = F$ where

$$G(N) = \frac{r\lambda_A}{(r + N\lambda_A)^2}\left(\pi_A + \frac{N\lambda_B}{r + N\lambda_B}V_A\right) + \frac{N\lambda_A}{r + N\lambda_A}\frac{r\lambda_B}{(r + N\lambda_B)^2}V_A.$$

Denote the solution to this equation N^*, and note that $G(N)N \to 0$ as $N \to \infty$.

Suppose that N firms have entered. In the subsequent game, the firm i splits its capacity between A and B according to $x_{i,A}$ and $x_{i,B}$ respectively. Rival firms split their capacity in such a way that there is an aggregate effort towards invention $j \in \{A, B\}$ (including that of the small firm) is z_j. After the first invention is discovered, the N firms will direct their

capacity towards the remaining invention. Ignoring the integer constraint, the zero profit condition is $H(N) = F$ where

$$H(N) = \sum_{j \in \{A,B\}} \frac{x_{i,j}\lambda_A}{r + z_A\lambda_A + z_B\lambda_B} \left(\pi_{1,j} + \frac{N\lambda_{-j}}{r + N\lambda_{-j}} \pi_{2,-j} \right) = F$$

Denote the solution to this equation N^e. Given that $x_{i,j} \leq 1$ and $z_j \leq N$, we have $H(N)N \to \Omega$ as $N \to \infty$, with $\Omega > 0$. This shows that, as $N \to \infty$, there will be a threshold $\bar{N}$ such that $H(N) > G(N)$ for all $N \geq \bar{N}$. As the severity of the crises increases and both N^* and N^e are above $\bar{N}$, we will have $H(N^e) = F = G(N^*) < H(N^*)$. If we select a type of equilibria with a particular feature (e.g., equilibrium where all firms work on a particular invention), then $H(\cdot)$ is decreasing, which implies that $N^e > N^*$.