Government Guided Funds and the Direction of Innovation*

Zhao Jin[1] Erica X.N. Li[2] Dan Su[3] Daniel Yi Xu[4]

May, 2026

Preliminary Draft. Please do not distribute.

Abstract

We study Chinese central government guided funds (GGF), which invest through private venture capital partnerships while pursuing industrial policy objectives. We argue that, beyond the standard knowledge externality, private investors under-invest in discovering which technological directions are viable because the resulting information spills over to competing firms. By financing directional experimentation, GGF could help resolve technological uncertainty and guides entrepreneurial activity toward viable technologies. Using a shift-share instrument that combines national industry-level fundraising with predetermined provincial industry shares, we find that central GGF investment crowds in private venture capital, increases firm entry, and raises the share of R&D-intensive firms in targeted industry-provinces. Patent portfolios become more similar over time, consistent with convergence toward viable technological directions. In contrast, local GGFs generate weaker and less persistent effects, suggesting that geographically constrained industrial policies are less effective at resolving technological uncertainty.

*We would like to thank Mariassunta Giannetti, Bing Han, Abhiroop Mukherjee, Jun Pan, Neng Wang, Dong Yan, Alminas Zaldokas for helpful discussions. All errors are our own.

[1] jinzhao@ckgsb.edu.cn, Cheung Kong Graduate School of Business.

[2] xnli@ckgsb.edu.cn, Cheung Kong Graduate School of Business.

[3] dansu@ckgsb.edu.cn, Cheung Kong Graduate School of Business.

[4] daniel.xu@duke.edu, Duke Univeristy & NBER.

1 Introduction

Government venture capital (GVC) has become an increasingly important vehicle for implementing industrial policy.[1] Despite its rapid growth, the economic rationale for GVC remains underexplored. Does government venture capital simply reallocate financing toward politically favored firms, or can it address market failures that private investors leave unresolved? We study China's Government Guided Funds (GGFs), one of the largest GVC programs in the world. The GGF refers to a fund established by either central or local government and operated in a market-oriented manner. GGFs invest in private VC funds as LPs. GGFs require lower or, in some cases, zero financial returns, but they impose constraints on the industries of investment to serve the government's industrial policies.

Several arguments have been advanced in favor of GGF investments. First, relative to direct subsidies and tax incentives, GGFs may represent a more efficient use of public fiscal resources. Unlike one-time transfers, equity investments can generate financial returns, allowing capital to be recycled and redeployed to support additional firms, particularly those in their early stages of development or facing financing constraints. Second and more importantly, emerging industries, due to their high R&D intensity, uncertain technological direction, and long investment return cycles, are often under-invested by private funds (e.g., Griliches, 1991; Jones and Williams, 1998; Hall, 2002). To the extent that GGFs internalize innovation spillovers and adopt longer investment horizons than private venture capital (PVC) funds, they may help alleviate these market failures. At the same time, government intervention is frequently associated with concerns about inefficiency and political distortions. Critics argue that GGFs may allocate capital to politically favored rather than economically efficient firms, leading to resource misallocation and reduced productivity (Acemoglu and Verdier, 2000).

We formalize these ideas in Section 3 in which a GGF and a counterpart private equity fund

[1] Examples include public equity-investment and fund-of-funds programs such as the European Innovation Council Fund, Bpifrance's direct and fund-of-funds venture investments, Germany's Zukunftsfonds/Future Fund, and the British Business Bank's Enterprise Capital Funds program. These programs use public capital to expand venture financing for innovative firms, often by investing directly in startups or indirectly through private venture funds.

are endowed with the *identical* per-firm cost, equity stake, screening capability, and enhancement technology; the only primitive that distinguishes the two institutions is the GGF's policy mandate, which causes it to internalize a fraction of the social returns from innovation. Within this framework, the economy suffers from two distinct market failures. The first is the familiar knowledge externality: innovators cannot fully capture the social benefits generated by their discoveries, leading to insufficient innovative activity. The second, more novel one, is what we call a *directional externality*. Before innovation can occur, investors must learn which technological directions are commercially viable. Because such assessment is costly and its benefits spill over to many firms within the industry, no individual entrepreneur has sufficient incentive to undertake it. As a result, promising technological opportunities may remain unexplored.

Traditional instruments such as subsidies and tax incentives can encourage firms to innovate once promising opportunities have been identified, thereby mitigating the knowledge externality. However, they do little to encourage the costly process of discovering which technological directions are worth pursuing in the first place. By contrast, a GGF can address both margins simultaneously: it encourages firm entry and innovation while also financing the industry-level exploration needed to identify viable technological opportunities.

The model draws three key empirical predictions. First, the GGF lowers its effective screening threshold and supports high-ability marginal entrepreneurs whom PE alone would not back, producing crowd-in of private capital. Second, patent similarity is predicted to be *non-monotonic*: flat or declining in the short run during the exploration phase, and rising in the medium run once GGF-backed successes reveal the viable direction to the broader market. Third, GGF effectiveness rises with the local talent pool and entrepreneurial base and falls with frictions that erode enhancement capability or raise screening costs.

To test these model predictions, we construct a shift-share instrument that interacts the annual total amount of capital raised by central GGF targeting specific industries (the "shift") with the predetermined provincial share of sales in each industry as of 2007 (the "share"). Our identification relies on the idea that variation in central GGF fundraising over time reflects shifts in nationwide

policy priorities, rather than responses to localized investment conditions. Central GGFs are established by national ministries and centrally controlled SOEs to implement broad, national industrial strategies, and their fundraising decisions are made at the central level, largely independent of province-industry-specific shocks.[2] Our estimates show that central GGF investments generate economically meaningful effects that unfold over time. Starting from the sample mean of central GGF investment (2.4 firms per industry-province-year), one additional investment is associated with a 27% increase in private VC investment the following year and a 14% increase by $t+3$, both relative to the sample-mean number of privately funded firms. The impact extends to entrepreneurial entry, which rises by 16% three years after the central GGF investment and by 21% at $t+5$. Over the longer horizon, industrial composition shifts toward innovation: the share of firms whose R&D intensity ranks in the top decile of their industry rises by 8 percentage points at $t+5$ and remains elevated at $t+7$.

Building on these baseline findings, we investigate the reduction of technological uncertainty as a key channel through which central GGFs catalyze private investment. We measure technological uncertainty using patent similarityspecifically, the cosine similarity of patent portfolios across firm pairs within the same industry, following Bloom, Schankerman, and Van Reenen (2013). Higher patent similarity indicates that firms are pursuing more convergent technological paths, suggesting lower technological uncertainty. The data exhibit a non-monotonic dynamic that aligns with our theoretical predictions: in the short run, central GGF investment is associated with flat or marginally negative patent similarity (the exploration phase), consistent with the entry-expansion effect drawing additional entrants who sample directions across the technological space. In the medium run, however, GGF-backed successes accumulate and the viable technological direction becomes inferable from realized outcomes; patent similarity rises sharply and is positive and statistically significant from $t+3$ through $t+5$ in IV specifications (the convergence phase).

In addition to central GGFs, provincial and municipal governments also establish local GGFs

[2]We corroborate this interpretation by examining the 2012 Strategic Emerging Industries (SEI) Initiative, whose ex-ante sectoral characteristics (high R&D intensity, technological lag, geopolitical sensitivity) indicate selection on long-run policy considerations rather than contemporaneous market performance.

to implement industrial policies. Theoretically, this decentralization presents a trade-off: while local governments may possess an informational advantage regarding firms within their jurisdictions, they are simultaneously restricted to a geographically constrained investment set. Our empirical results suggest that these constraints limit the efficacy of local funds relative to their central counterparts: whereas central GGFs drive persistent innovation, local GGFs generate only short-lived increases in capital and entry without significant technological spillovers.

These findings contextualize the work of Guo and Zhao (2025), who similarly document positive associations between local GGFs and private participation. We argue, however, that interpreting such average effects is difficult given that local industrial policy is endogenous–targeted industries are likely selected based on unobserved potential. We find that local GGFs effectively crowd in private capital and entrepreneurs only in regions with robust economic endowments and high-quality governance. This heterogeneity highlights the critical role of local context in shaping the efficacy of government industrial policy: while central GGFs can catalyze private investment across diverse settings, local GGFs appear to require a conducive environment to successfully crowd in private capital and foster sustainable industry growth.

Our paper contributes to the field of industrial policy, a topic that has been explored since the seminal works of Rosenstein-Rodan (1943) and Hirschman (1958). Historically, the prevailing view was that industrial policies lead to inefficient resource allocation and adversely affect developing economies (Krueger, 1990). However, recent research has begun to explore the underlying rationale and potential positive impacts of industrial policies. For instance, Aghion et al. (2015) demonstrate that Chinese industrial policy has increased productivity growth by promoting competition. Besides, Itskhoki and Moll (2019) analyze the optimal dynamic Ramsey policies and conclude that wages-suppressing industrial policies can mitigate the distortions caused by financial frictions. Moreover, Liu (2019) suggests that a well-intended government may have reasons to subsidize upstream sectors, which could lead to positive overall economic effects. Additionally, Liu and Ma (2023) examine the trade-offs between allocating more R&D resources to central sectors in the innovation network and leveraging foreign knowledge spillovers. Finally, Ju et al.

(2024) argue that the "Made-in-China 2025" industrial policy tends to improve the welfare of both China and the U.S. Closest to our setting, Guo and Zhao (2025) study local government guidance funds. We complement their analysis by focusing on centrally guided funds, for which our shift-share design delivers cleaner causal identification, while inference for local funds is complicated by potentially endogenous industry targeting.

The paper is organized as follows. Section 2 provides institutional background and details on the data. Section 3 develops a theoretical framework that explains how GGF promotes industrial development. Section 4 presents the shift-share instrumental variable approach and the main empirical results on the effects of central GGF investments. Section 5 explores the reduction of technological uncertainty as a key mechanism behind the crowding-in effects. Section 6 examines the heterogeneity of local GGF performance across different contexts. Section 7 concludes.

2 Institutional Background, Data, and Sample Construction

In this section, we first provide institutional details of China's two-tired government guided fund and outline its pros and cons relative to other industrial policy tools such as subsidies and tax exemptions. We then describe the data sources and empirical procedure used to construct our sample. Finally, we present the summary statistics of the sample.

2.1 Institutional Background

During the past two decades, Chinese governments at various levels have strategically invested in early-stage technology startups across key industries, aiming to advance national industrial policy objectives. Government Guided Funds (GGFs) serve as the primary mechanism for channeling public capital into the venture capital (VC) ecosystem. These funds are established by state-owned enterprises, central government ministries, or local administrations, with a structured approach to fostering technological innovation and industrial upgrading.

While a small fraction of GGFs directly invest in companies, the majority operate as limited

partners (LPs) in private VC/PE funds.[3] Typically, GGFs contribute 5% to 30% of the capital for each fund, leveraging the market discipline and professional expertise of private-sector fund managers. In this paper, we refer to VC/PE funds receiving GGF capital as "Government VCs" (GGFs), and funds receiving 100% private capital as "private VCs" (PVCs). Compared to conventional tools like subsidies and tax exemptions, GGFs offer distinct structural advantages that address long-standing inefficiencies in industrial policy.

Targeted Allocation of Resources. Unlike blanket subsidies that distribute funds across all firms in target industries, GGFs focus on identifying and funding the most promising startups by delegating investment decisions to professional fund managers, given that governments often lack the expertise to distinguish between viable and non-viable ventures. That is, GGFs mitigate the information asymmetry inherent in traditional policies. This mechanism ensures that public funds are directed toward high-potential enterprises, enhancing capital efficiency.

Ex-ante Commitment Mechanism. GGF investments serve as a pre-emptive commitment from governments to support targeted industries, contrasting with subsidies and tax exemptions, which are typically offered as ex-post rewards to firms that have already entered the sector. The uncertainty around whether ex-post support will be fulfilled creates hesitation among investors and entrepreneurs. GGFs, by contrast, provide clear upfront signals of government commitment, reducing investment risks and encouraging private-sector participation.

Temporary and Self-Sustaining Funding Model. GGFs adopt a "patient capital" approach, investing in early-stage companies and exiting once they reach maturity either through acquisition/IPO or transferring shares to private investors. Unlike subsidies and tax exemptions, which often distort market selection by keeping inefficient firms alive through ongoing policy support, GGF funding is merely a temporary bridge to financial independence. This model not only avoids long-term fiscal burdens but can generate returns: successful exits via IPOs or acquisitions may yield fiscal revenues, transforming GGFs from cost centers into potential sources of public income.

These advantages have led to a rapid expansion in both the number of GGFs and their share of

[3]Only 4% are managed by state-controlled VC firms, as noted by Fei et al. 2018

the VC/PE market over the past two decades. In Figure 2, we present the total number of central GGFs and the proportion of GGFs by fund establishment year from 2007 to 2019. In 2007, 53 GGFs were established. The annual growth rate of GGFs has consistently been positive with an average of about 42%. The peak of GGF's expansion occurred in 2017, with a record of 2,684 GGFs established. However, following this peak year, the number of GGFs began to decline. In 2019, 1,772 GGFs had been established. The proportion of GGFs relative to all VC/PE funds in China stays stable around 15% with a notable spike to over 20% in 2019.

In Figure 3, we plot the number of firms invested by GGFs and by PVC funds solely separately. The overall pattern is consistent with the number of funds established in each year in Figure 2. One notable difference is that the proportion of number of GGFs relative to all the VC/PE funds is about 15%, while the number of firms that are invested by GGFs constitutes about 53%.

In sum, Government Guided Funds represent a new approach to industrial policy, combining public strategic objectives with private-sector operational efficiency. By addressing information asymmetry, strengthening policy commitment, and adopting a self-limiting funding model, GGFs potentially offer a path toward more cost-effective and sustainable public investment. However, they are not immune to the risks of government intervention in markets. Funding may flow to VC/PE funds with political connections rather than genuine investment acumen. And government agencies may interfere in GGFs' investment decisions, compromising professional autonomy. Whether—and under what scenarios—GGFs generate positive or negative net effects in China remains an open empirical question, which the analysis below seeks to address.

2.2 Data and Sample Construction

2.2.1 Government Guided Funds

To identity GGFs, we first collect the universe of 15,297 VC/PE firms from Asset Management Association of China (AMAC). We then search these VC/PE firms in the business registration database maintained by the State Administration of Industry and Commerce of China (SAIC). Of 15,297 VC/PE firms, we were able to find 14,214 firms by searching firm names. We then identify

84,920 funds established from 2007 to 2019 by those 14,214 VC/PE firms.

The database provides shareholders of each legal business entity in China. That is, for each VC/PE fund, we can observe its LPs, the shareholders of each LP, and the shareholders of each shareholder. We collect four layers of shareholders of each LP. Moreover, we also observe the companies invested by these VC/PE funds.

To identify whether a LP has government ownership, we search whether the four-layers of shareholders are state-owned enterprises (SOEs), central, or local government agencies. To identify government agencies, we first use the shareholder classification provided by SAIC that indicates whether a shareholder is a government agency. In instances where shareholder classifications are absent, we compile an additional list of shareholders by searching for specific keywords such as *government*, *bureau*, *department*, and *administration* within their names. Subsequently, we manually verify whether these entities indeed represent government agencies. Furthermore, to distinguish between local and central government agencies, we manually classify them based on names. To identify SOEs, we collect the names of a comprehensive list of SOEs from the State-owned Assets Supervision and Administration Commission, which we match to the business registration data. We then manually classify whether a SOE is supervised by a central or local government by searching available information on internet.

2.2.2 Capital raised by Central GGFs and their target industries

To construct our instrumental variable shift-share, we hand-collect the total amount of capital raised by central GGFs and their target industries. Information on these funds is collected from public sources, including government announcements, news coverage, and fund websites. Funds lacking information on designated industries or total capital raised are excluded. Our final dataset includes 350 funds established between 2009 and 2016 by central government ministries or centrally controlled SOEs. For funds targeting multiple industries, we apportion the total capital equally across all targeted industries. On average, these funds raised RMB 1.4 billion (approximately USD 200 million) and targeted about two industries.

2.2.3 New firm entry

We collect the new firm entry data from the business registration database maintained by SAIC that we used to identify GGFs. We then aggregate the new firm entries at province-industry level. The industry classification provided by SAIC is based on 20 sectors, and we exclude four industries in our analyses: 1) financial services, 2) public management, social security, and social organizations, 3) international organizations and foreign institutions, and 4) water conservancy, environment, and public facilities management (utility).[4]

By combining databases from various data sources, the final province-industry sample consists of 3,956 observations spanning 31 provincial-level regions in Mainland China and 16 industries from 2006 to 2016.[5] The industry classification is provided by the business registration database (SAIC).

2.3 Summary statistics

We present two sets of summary statistics. Table 1 provides summary statistics for the data at province-industry level in Panel A and the data at firm level in Panel B. At the province-industry level, local GGFs invest in 2.52 firms, central GGFs invest in 2.41 firms, and PVCs invest in 2.73 firms on average. The distributions of investments at province-industry level are highly right-skewed for both GGFs and PVCs, indicating a highly selective investment pattern focused on a small number of industries. In addition, the average number of new firm entry is 116.17 and is also right-skewed with median being 26.

[4]The 16 sectors included in our analyses are: agriculture, forestry, animal husbandry, and fishery; mining and quarrying; manufacturing; production and supply of electricity, heat, gas, and water; construction; wholesale and retail trade; transportation, warehousing, and postal services; accommodation and catering services; information transmission, software, and information technology services; real estate; leasing and business services; scientific research and technical services; water conservancy, environment, and public facilities management; resident services, repair, and other services; education; health and social work; culture, sports, and entertainment.

[5]After the outbreak of the U.S.-China Trade War in 2018, private venture capital and private equity activity in China contracted sharply, with foreign-backed and purely private investors significantly reducing commitments amid heightened geopolitical uncertainty and regulatory scrutiny. Empirical data indicate a marked pullback in U.S. and other private capital flows into Chinese VC/PE markets in the years following 2018. As a result, a growing share of remaining funds now incorporate government capital or are directly government-linked, reducing the availability of truly private VC/PE in the post-2016 period and rendering data after 2016 less suitable for our empirical comparison of GGFs and PVCs.

Firm-level sample consists of 92,833 observations with 10,155 unique firms from 2007 to 2016. The sample includes firms that are invested by either GGF or PVC at least once during the period of 2007 to 2016. The average of dummy variable *Central_GGF* is 0.12, indicating that 12% of firm-years involving central GGFs' investments. Likewise, 13% of firm-years involving local GGFs' investments, as shown by the average of *Local_GGF*, and 14% of firm-years involving PVCs' investments. The average number of granted patents that are applied no later than a given year is 1.85, with a median value of zero.

Table 2 presents summary statistics on the characteristics of firms receiving different forms of government support. Panel A of the table reports the average firm characteristics across three groups: firms that receive only government subsidies (Sub_only), firms that receive only Government Venture Capital (GGF_only), and firms that receive both subsidies and GGF investment (Sub_GGF). Panel B conducts the comparison via OLS regression with industry-year fixed effects. The set of firm characteristics include return on assets excluding subsidies (ROA_ExL_Subsidy), sales, total assets, average wage, R&D activity (including non-zero R&D, R&D-to-sales ratio), patent activity (including the probability of holding any patents, Non-zero Patent, and the number of patents, Num_Patent), and firm age.

Table 2 suggest that GGF investment is associated with significantly higher firm sales, total assets, and average wages, indicating GGF-backed firms are larger and potentially more productive. However, GGF investment alone does not significantly affect R&D intensity or the likelihood of conducting R&D. GGF-backed firms are also significantly younger.

The interaction term GGF$\times$Subsidy reveals that firms receiving both forms of support show significantly higher R&D engagement compared with firms receiving GGF investments only: they are 27 percentage points more likely to conduct R&D and show a 1.4 percentage point increase in R&D-to-sales ratio. They are also more likely to hold patents (3.8 percentage points) and exhibit significant improvements in ROA. These findings suggest that while GGFs drive up firm scale and innovation outcomes, their impact on R&D activity is magnified when paired with direct government subsidies. This highlights the importance of a coordinated policy approach that combines

capital provision with direct incentives for innovation.

3 GGF Investments and Industrial Development

This section develops a theoretical framework for how GGF promote industrial development. Two market failures lie at its center, and we argue that conventional policy instruments cannot resolve both of them. In the model, entrepreneurs make two endogenous decisions: which technological direction to pursue and whether to enter. The GGF plays three roles: it screens high-ability entrepreneurs, enhances their innovation capability, and conducts directional research that guides portfolio firms toward viable technological paths. We give the GGF the same per-firm cost, equity stake, and screening and enhancement technology as a counterpart private equity fund. This is a deliberately demanding benchmark. It ensures that our conclusions about why GGF crowds in private capital, rather than substituting for it, rest on the GGF's policy mandate alone, not on any assumed cost or financing advantage.

We keep technical material out of the main text. The formal statements of assumptions, closed-form expressions, first-order conditions, and proofs are collected in Internet Appendix C. Here we lay out the environment, the two market failures, the comparison between GGF and other policy instruments, and the empirical predictions. We state the lemmas and propositions but discuss them through their economic interpretation.

3.1 Model Framework

The economy lasts three periods and consists of a traditional sector and a modern sector. In the traditional sector, technologies are known ex ante and labor exhibits diminishing returns. In the modern sector, productivity is initially unknown and can only be discovered through costly innovation. Labor supply is fixed, and the wage clears the labor market each period.

In period 0, entrepreneurs with privately known ability $z \in [\underline{z}, \bar{z}]$ decide whether to pay a fixed entry cost f to explore a modern good. In period 1, successful innovators observe productivity and

produce under temporary monopoly of length T. In period 2, productivity information diffuses, entry becomes unrestricted, and the economy converges to a competitive equilibrium.

Direction choice versus productivity draw. The modern sector consists of a continuum of goods, each belonging to one of $K \geq 2$ *technological directions* $d \in \{1, 2, \ldots, K\}$, representing distinct technological paths (e.g., alternative battery chemistries in the EV industry or different semiconductor architectures). One direction d^* is viable, and it leads to high long-run productivity. However, d^* is unknown ex ante to all agents.[6]

The model distinguishes sharply between two endogenous objects, and this distinction is central to the mechanism:

- *Direction choice* d_i. Each entering entrepreneur *chooses* which technological direction to pursue, $d_i \in \{1, \ldots, K\}$. Empirically, the direction choice corresponds to the firm's R&D agenda, that is, the technological class in which it concentrates its research. It is the counterpart of the patent-class composition of a firm's patent portfolio.

- *Productivity draw* θ_i. Conditional on the chosen direction d_i and the entrepreneur's ability z_i, productivity $\theta_i \in [0,1]$ is then *drawn* from a power distribution $F^*(\cdot; z_i)$ if $d_i = d^*$ and $F^0(\cdot; z_i)$ otherwise, with F^* first-order stochastically dominating F^0 at every z. The draw represents the realized success of the firm's R&D effort.

Without information about d^*, an entrepreneur chooses each direction with equal probability $1/K$. The closed-form distributions and parameter restrictions are in Appendix C.1.

Directional research and information diffusion. Before entrepreneurs commit to a direction, any agent who controls a portfolio of firms (either an investor, a fund, or a planner) can spend resources $\kappa \geq 0$ on technological assessment, such as industry analysis, pilot projects, and technology scouting. This effort reveals d^* with probability $q(\kappa)$, where q is strictly increasing and

[6] For tractability, we assume a single viable direction. The qualitative results extend to the case where a strict subset of directions is viable, with d^* reinterpreted as the set of high-productivity paths.

strictly concave with $q(0) = 0$ and $0 < q'(0) < \infty$. An agent who learns d^* shares it only with its own portfolio firms. Firms outside the portfolio keep choosing directions uniformly. Two features of directional research, both formalized in the appendix, are central to the model:

- *Indivisibility.* Directional research is a portfolio-level fixed cost in substance: its benefit scales with the number of firms directed, while any single entrepreneur benefits through only one draw. Hence it is uneconomical for individual entrepreneurs who control a single firm.

- *Portfolio scale.* For any fund or planner operating across many firms, the marginal portfolio value of directional research at $\kappa = 0$ exceeds its marginal cost, so each portfolio problem has an interior solution for κ.

In other words, industry-wide technological assessment is not worth much to a single entrepreneur, but it becomes valuable once its signal can be spread across a portfolio. The atomistic outcome $\kappa^d = 0$ therefore does not mean that technological information has zero *private* value when firms are bundled. It reflects the absence of any agent who internalizes the portfolio-level return.

After period-1 production, the realized productivities and the directions chosen by all entrants become public. Agents then form beliefs about d^* by maximum likelihood, inferring d^* to be the direction with the highest realized productivity at the top. In period 2, new entrants and pivoting firms converge on this inferred direction. The diffusion channel thus runs through outcomes: directional research, once GGF-backed firms turn it into successful products, spreads to the rest of the market.

Productivity frontier and patent similarity. Let m denote the number of period-1 entrants and $D^* \subseteq \{1, \ldots, m\}$ the (random) set who pursue d^*. The economy's productivity frontier is $\hat{\theta} \equiv \max_{i=1,\ldots,m} \theta_i$, which determines long-run productivity through period-2 aggregate output. The next lemma summarizes its key properties.

Lemma 1 (Expected Innovation Frontier)**.** *Conditional on the directional allocation D^* and ability profile $\{z_i\}$, the expected productivity frontier $\mathbb{E}[\hat{\theta} \mid D^*]$ is strictly increasing in (a) the number of*

entrants m (extensive margin), (b) the abilities $\{z_i\}$ (intensive margin), and (c) the share and quality of entrepreneurs allocated to d^* (directional margin). The marginal value of reallocating an entrepreneur from a wrong direction to d^* is strictly increasing in that entrepreneur's ability z.

The lemma identifies three margins along which the frontier moves, with its technical form given in the appendix. The frontier depends both on how many entrepreneurs innovate and on how their ability is distributed across technological paths. High-ability entrepreneurs contribute disproportionately, and they contribute even more when steered toward d^*. This convexity of the directional gain in ability is the basic force behind the value of combining GGF screening with directional guidance.

We define patent similarity $\mathscr{S}$ as the cross-direction concentration of *direction choices*:

$$\mathscr{S} \equiv \sum_{k=1}^{K} \left(\frac{n_k}{m}\right)^2, \tag{1}$$

where n_k is the number of firms with $d_i = k$. Under uniform allocation, $\mathbb{E}[\mathscr{S}] \to 1/K$ as $m \to \infty$, while under convergence to a single direction $\mathscr{S} \to 1$. By construction, $\mathscr{S}$ depends on the choice margin $\{d_i\}$ rather than on the productivity draws. This is why the model speaks directly to the dynamics of patent-portfolio concentration, the measure of technological uncertainty used in the empirical literature.

On the demand side, each good has unit demand at willingness to pay p, firms produce in period 1 if their unit cost does not exceed p and earn monopoly profits for duration T, and in period 2 entry becomes unrestricted and competition eliminates monopoly rents.

3.2 Two Market Failures

With the environment fixed, we ask how the decentralized equilibrium compares to the social optimum. The planner has two instruments, the number of entrants m and directional research κ, and chooses them to maximize total surplus, trading off the period-0 traditional output forgone against the period-2 frontier improvement. This yields two first-order conditions. The *scale* condition

equates the marginal cost of entry to its marginal social benefit through the frontier. The *direction* condition equates the marginal cost of directional research to its marginal social benefit through better directional allocation. The planner needs two instruments because there are two distinct sources of social value. The formal problem and its first-order conditions are in the appendix.

In the decentralized economy, atomistic entrepreneurs enter when expected monopoly profit exceeds the entry cost, yielding equilibrium entry m^d. Because no entrepreneur controls a portfolio, directional research is zero by Indivisibility: $\kappa^d = 0$ and hence $q^d = 0$. Comparing the decentralized outcome to the planner's optimum yields:

Proposition 1 (Two Market Failures)**.** *The decentralized equilibrium is inefficient on two distinct margins relative to the social optimum:*

(i) ***Knowledge externality.*** *There exists a unique threshold $\hat{T} < \infty$ such that for any $T < \hat{T}$, $m^d < m^*$.*

(ii) ***Directional externality.*** *In the atomistic decentralized equilibrium, $\kappa^d = 0 < \kappa^*$.*

The two wedges are conceptually distinct, even though both arise from the public-good nature of innovation outcomes. The knowledge externality is the standard innovation-spillover wedge. Private inventors capture only the period-1 monopoly rents, which are proportional to T, whereas the planner also values the period-2 frontier improvement that eventually diffuses to consumers and to imitating firms. The directional externality is separate. Even if knowledge spillovers were fully internalized, no single entrepreneur would pay for industry-wide assessment, because each captures only the gain from one draw while bearing the full cost of a non-rival signal. The inequality $\kappa^d = 0 < \kappa^*$ thus compares the atomistic benchmark with the social value of information applied across the whole economy. It does not rule out positive directional research by a portfolio investor, which we introduce next. Because these are two independent wedges rather than one wedge with two symptoms, a useful policy instrument must address both.

3.3 GGF versus Private Equity: The Role of Mandate

With the two market failures in hand, we compare the GGF against two private alternatives, direct entrepreneurship (no fund) and a private equity (PE) fund, and against three traditional policy instruments (patents, subsidies, and tax incentives). The traditional instruments address the knowledge externality by raising entry incentives, but none addresses the directional externality: they convey no information about d^* and do not screen for entrepreneur quality. We begin with the comparison between GGF and PE, which is the most demanding one, since PE is itself an intermediated investor with screening and enhancement capabilities. We defer the traditional instruments to Section 3.5.

We adopt the most demanding setup for evaluating the GGF. The GGF and its PE counterpart share the same per-firm screening cost $c > 0$, the same equity stake $\gamma \in (0,1)$, the same directional-research technology $q(\cdot)$, and the same enhancement technology $\bar{b} > b$. The GGF therefore has no informational advantage, no cheaper capital, and no lower screening cost than PE.

The two institutions differ along one dimension only. The GGF internalizes a fraction $\lambda \in (0,1)$ of the social returns from frontier improvement and information diffusion, while the PE fund maximizes only its own equity returns ($\lambda = 0$). We treat $\lambda > 0$ as a primitive of the GGF's institutional design rather than something derived. GGF charters require the fund to support strategic industries, government LPs accept below-market IRRs, and policy guidelines tie GGF managers' performance to industrial-development objectives.[7] When $\lambda = 0$ the GGF coincides exactly with a PE fund, so in equilibrium the entire wedge between the two institutions is the mandate λ.

The GGF's decisions proceed in two stages. In Stage 1, taking the supported set H as given, the GGF chooses how much directional research κ^{GGF} to conduct, equating its marginal value to the supported portfolio (a private-return term plus a social-return term weighted by λ) to its marginal cost. In Stage 2, given the period-1 directional outcome from Stage 1, the GGF supports

[7] Survey evidence in Colonnelli, Li, and Liu (2024) confirms that government LPs in Chinese venture-capital partnerships routinely accept lower returns in exchange for industrial-policy targeting, and the State Council documents underlying the Strategic Emerging Industries Initiative tie GGF activity to long-run social returns rather than financial performance.

entrepreneur i if and only if the equity-weighted private gain plus the λ-weighted social gain exceeds the per-firm cost c. The PE counterpart solves the same problem with $\lambda = 0$. The Stage 1 and Stage 2 first-order conditions are in the appendix.

The comparison between GGF and PE reduces to comparing two screening rules that differ *only* by a $\lambda \Delta W(z)$ term, where $\Delta W(z)$ is the social-welfare gain from supporting an entrepreneur of ability z. Under a single-crossing condition on the private and social gains, this yields:

Lemma 2 (Screening Threshold)**.** *For any fixed directional outcome $q \geq 0$, there exist unique screening thresholds $\tilde{z}^{GGF}(q,\lambda)$ and $\tilde{z}^{PE}(q)$. Moreover, $\tilde{z}^{GGF}(q,\lambda) < \tilde{z}^{PE}(q)$ if and only if $\lambda > 0$.*

In effect, the GGF discounts the screening cost by $\lambda \Delta W(z)$, which turns a marginal entrepreneur with low private profitability but high social value into a supportable target. The "if and only if" part pins down the role of internalizing the social value: when $\lambda = 0$ the GGF degenerates into a PE fund and the two thresholds coincide, and the strict ordering appears only when the mandate is active. Since ΔW is increasing in z, the gap between the two thresholds is largest where the social return is largest, namely for high-ability entrepreneurs whose success generates large general-equilibrium spillovers.

Corollary 1 (Marginal Entrepreneurs Backed by GGF Only)**.** *Entrepreneurs with ability $z \in [\tilde{z}^{GGF}, \tilde{z}^{PE})$ are supported by the GGF but* not *by any PE fund operating with the same (c,γ). The mass of such entrepreneurs is the GGF's "additional" coverage relative to PE, and is strictly positive whenever $\lambda > 0$ and the ability distribution is non-degenerate on this interval.*

Corollary 1 clarifies where the additional GGF activity comes from: GGF-backed firms are not just firms that PE would have backed anyway. With the same (c,γ) technology, the GGF reaches a strictly larger pool, and the extra coverage is concentrated on entrepreneurs whose private profitability alone does not justify support. These entrepreneurs, in $[\tilde{z}^{GGF}, \tilde{z}^{PE})$, form the GGF's mandate-only portfolio. They are the firms that any PE fund would leave unbacked at the prevailing state of the industry, but that the GGF supports because its mandate $\lambda > 0$ lets the social return cover the private-NPV shortfall.

Joint equilibrium of screening and directional research. Stages 1 and 2 are linked: the threshold sets the portfolio size, the portfolio size sets the marginal value of κ, κ determines q, and q feeds back into the threshold. Under a stability condition on this joint problem, the GGF equilibrium $(\kappa^{GGF}, \tilde{z}^{GGF})$ is unique, and the analogous PE problem yields a unique $(\kappa^{PE}, \tilde{z}^{PE})$. The appendix shows that the composite best-response map has a unique interior fixed point with local slope below one. The next lemma states the resulting ordering.

Lemma 3 (Joint GGF Equilibrium)**.** *At the stable joint equilibria,*

$$0 = \kappa^{d} < \kappa^{PE} < \kappa^{GGF} < \kappa^{*} \quad \textit{and} \quad \tilde{z}^{GGF} < \tilde{z}^{PE}, \quad |H^{GGF}| > |H^{PE}|.$$

The mechanism is one of reinforcement. Under the GGF, a lower threshold expands the portfolio and raises the marginal value of directional research, because more firms benefit from each unit of κ. This pushes κ^{GGF} above κ^{PE}. The higher q^{GGF} in turn raises the private and social gains for any given entrepreneur, which lowers the threshold further. PE has no λ wedge to set off this loop, so its joint equilibrium settles at strictly less screening and strictly less directional research. The planner does more still on both margins, because it values the entire social return rather than only the fraction $\gamma + \lambda < 1$ that the GGF captures.

Why does private capital alone fail? Because the setup is symmetric, the answer must rest on $\lambda > 0$ rather than on any cost or technology advantage. The two market failures of Proposition 1 are coordination failures from the viewpoint of any single agent. When a fund supports a high-ability entrepreneur, the resulting frontier improvement benefits the whole sector once information diffuses, so each fund free-rides on others' screening and too few entrepreneurs are supported. When a fund's portfolio firm reveals d^*, the inference rule turns that information into a public good: the fund bears the full cost κ of directional research but captures only the equity share γ of its own portfolio's gain, not the sector-wide benefit of revealing d^*. A PE fund with the same $(c, \gamma, q, \bar{b})$ as the GGF internalizes none of this public-good component, and by Lemma 3 it sets $\tilde{z}^{PE} > \tilde{z}^{GGF}$ and $\kappa^{PE} < \kappa^{GGF}$. The GGF closes part of both gaps precisely because $\lambda > 0$. It

acts on the marginal pool $[\tilde{z}^{GGF}, \tilde{z}^{PE})$ that no PE fund backs in period 0, and it directs these firms toward d^*. The period-1 successes that follow raise the productivity distribution of later entrants and reveal d^* to the broader market through outcome inference.

3.4 Real Effects: Productivity and Convergence Dynamics

We now translate the joint equilibrium of Lemma 3 into implications for the distribution of producer productivities and the dynamics of patent similarity.

Proposition 2 (Quality Filtering)**.** *Let $\mathscr{F}_{GGF}(\theta)$ and $\mathscr{F}_{NoGGF}(\theta)$ be the CDFs of potential productivities among period-1 entrants before applying the endogenous production cutoff. Then $\mathscr{F}_{GGF}$ first-order stochastically dominates $\mathscr{F}_{NoGGF}$:*

$$\mathscr{F}_{GGF}(\theta) < \mathscr{F}_{NoGGF}(\theta) \quad \textit{for all } \theta \in (0,1).$$

Three reinforcing channels shift the entrant-productivity distribution rightward under the GGF: *selection* (the GGF supports the marginal entrepreneurs in $[\tilde{z}^{GGF}, \tilde{z}^{PE})$ from Corollary 1), *capability boost* (supported firms operate with the enhanced productivity exponent $\bar{b} > b$ on each direction, which FOSD-shifts their conditional productivity), and *directional guidance* (supported firms are disproportionately on d^*, which itself FOSD-dominates the other directions). We state the result for potential productivities before the production cutoff: with unit output per firm, labor demand per firm falls in productivity, so the sign of the equilibrium wage response is generally ambiguous. The quality implication of the model thus rests on the productivity distribution produced by screening, enhancement, and direction choice, not on an assumed wage increase. The CDF expressions and the proof are in Appendix C.3.5.

Proposition 3 (Technological Convergence)**.** *Let $\mathscr{S}_t^{GGF}$ and $\mathscr{S}_t^{NoGGF}$ denote patent similarity at time t.*

(i) Short run. *Under an explicit condition on the entry-expansion effect, $\mathbb{E}[\mathscr{S}_1^{GGF}] < \mathbb{E}[\mathscr{S}_1^{NoGGF}]$: the entry-expansion effect dominates the directional-concentration effect.*

(ii) Medium run. *Under the outcome-inference diffusion rule,* $\Pr(\hat{d} = d^* \mid GGF) > \Pr(\hat{d} = d^* \mid NoGGF)$*, and hence* $\mathbb{E}[\mathscr{S}_2^{GGF}] > \mathbb{E}[\mathscr{S}_2^{NoGGF}]$.

(iii) Monotonicity. $\partial \mathbb{E}[\mathscr{S}_2^{GGF}]/\partial \kappa^{GGF} > 0$.

The non-monotonic dynamic captures a basic trade-off in industrial policy. By lowering the screening threshold, the GGF pulls additional entrepreneurs into the modern sector. Some of them would otherwise have stayed in the traditional sector, and they now sample directions across the technological space, which can temporarily *reduce* concentration. The same intervention, however, steers high-ability supported firms onto d^*, where their high $\bar{b}$ delivers high realized productivities. As these successes accumulate, the diffusion mechanism reveals d^* to the broader market, new entrants and pivoting firms converge on d^*, and patent similarity rises. The GGF never discloses d^* to the whole market directly. It shares directional information only with its portfolio firms, and everyone else learns from observed outcomes. Convergence is therefore a byproduct of GGF-backed successes, not a deliberate policy of public disclosure.

3.5 Comparing Policy Instruments

The comparison with PE is one benchmark. The comparison with traditional fiscal instruments, such as subsidies, tax incentives, and patent extensions, is also important. To place all instruments on a common scale, we give each the same fiscal budget B and ask which dollars buy the largest welfare improvement. The detailed budget identities and welfare functionals are in Appendix C.2.

The feature that separates these instruments is structural. Subsidies, tax incentives, and patent extensions all operate on the individual entrepreneur's profit, whereas directional research is a portfolio-level fixed cost. An instrument that does not bundle entrepreneurs into a portfolio cannot deliver $q > 0$, however large its budget. Subsidies, tax incentives, and patent extensions are thus "one-knob" instruments: they can buy more entry but not better directional allocation. The GGF is the only "two-knob" instrument in the set, which moves both m and q as its budget grows. As a result, the welfare ranking is therefore conditional, not universal. When the directional exter-

nality is small or absent, as in a hypothetical industry where d^* is already known, the traditional instruments can outperform the GGF because they move m more cheaply per dollar. When the directional externality is large, as in emerging industries with high R&D intensity and uncertain technological paths, the GGF is welfare-maximizing, because the social value of the extra movement in q outweighs any loss on the extensive margin.

3.6 Determinants of GGF Effectiveness

So far the analysis shows whether the GGF works in principle. We now ask when it works best. The model points to several parameters whose comparative statics on the joint equilibrium $(\kappa^{GGF}, \tilde{z}^{GGF})$ are unambiguous, and these parameters map naturally onto observable institutional and regional characteristics.

Proposition 4 (Determinants of GGF Effectiveness)**.** *At the joint equilibrium, the expected frontier* $\mathbb{E}[\hat{\theta}^{GGF}]$ *satisfies:*

(i) ***Enhancement capability.*** $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}\bar{b} > 0$. *Higher enhancement raises the frontier and lowers* $\tilde{z}^{GGF}$.

(ii) ***Screening cost.*** $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}c < 0$. *Higher per-firm cost raises* $\tilde{z}^{GGF}$*, shrinking the supported set.*

(iii) ***Directional research technology.*** $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}q > 0$ *and* $\mathrm{d}\mathbb{E}[\mathscr{S}_2^{GGF}]/\mathrm{d}q > 0$. *A better directional-research technology shifts mass toward* d^**, raising the frontier and increasing medium-run patent similarity.*

(iv) ***Talent pool quality.*** $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}\bar{z} > 0$. *A richer talent distribution provides more high-ability candidates above the screening threshold.*

(v) ***Entrepreneurial pool size.*** $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}m > 0$. *A larger entrepreneurial base provides more firms on which the GGF's screening, enhancement, and directional channels can operate.*

The intuition for each comparative static is straightforward. Higher enhancement capability raises the frontier directly, because supported firms' productivity exponents rise, and by single-crossing it weakly expands the set of supportable entrepreneurs. A higher screening cost shrinks the supported set and mechanically reduces the GGF's reach. A better directional-research technology shifts mass toward d^* and so multiplies the impact of each supported firm. A richer talent pool or a larger entrepreneurial base gives all three GGF channels more to work with: more candidates above any given threshold, and more firms to direct at any given q. These are total comparative statics of the stable equilibrium. They follow from the threshold equation and the directional-research first-order condition, not from applying the envelope theorem to an object that is not the fund's value function. Full derivative expressions and the implicit-function arguments are in Appendix C.3.10.

These parameters map onto observable proxies for institutional and regional characteristics. The corruption index captures resource diversion, which lowers the GGF's effective enhancement capability $\bar{b}$. The bureaucratic-efficiency index captures administrative friction, which raises the effective screening cost c. Provincial GDP, the number of universities, and average wages proxy for the local entrepreneurial pool size m and the upper tail of the talent distribution $\bar{z}$. The GGF's geographic and sectoral scope governs the effective q: a fund spanning a wider sectoral and geographic horizon sees more pilot signals and a larger informational base for technological assessment, so its $q(\kappa)$ schedule lies above that of a narrowly scoped fund.

3.7 Empirical Predictions

The model delivers four empirical predictions. None is a direct restatement of a single proposition. Each instead combines several results into a hypothesis that is natural to frame in industry-province-time data. Together they motivate the empirical analysis below.

Prediction 1 (Crowd-In Effect of GGF)**.** *GGF intervention raises private VC investment in the same industry-region. The crowd-in works through two channels with different timing. The first is a productivity channel: higher demonstrated productivity raises expected returns to private investors*

soon after GGF entry. The second is an uncertainty-reduction channel: the viable direction d^* is revealed gradually through the high-productivity outcomes of GGF-backed firms, which lowers perceived risk. The productivity channel shows up at short horizons, whereas the uncertainty channel appears only once enough GGF-backed outcomes have accumulated.

By Lemma 3 and Proposition 2, the realized state under the GGF regime generates a rightward shift in the productivity distribution and a tighter posterior on d^*. Therefore, it is strictly more favorable than under the no-GGF regime. The underlying reason is that the GGF is the only fund that operates on the mandate-only pool $[\tilde{z}^{GGF}, \tilde{z}^{PE})$ from Corollary 1 and turns it into productive outcomes. Since PE's screening rule depends only on its own private NPV, the PE threshold $\tilde{z}^{PE}$ applied to the post-GGF state is strictly lower than the same threshold applied to the no-GGF state. We call this threshold-lowering effect of GGF presence on PE's screening *crowd-in*. We do not model a dynamic re-screening step explicitly, but the logic carries over from the static comparison: the GGF improves the state on which PE screens, and PE then finds a strictly larger set of entrepreneurs profitable.

Prediction 2 (Non-Monotonic Patent Similarity)**.** *Patent similarity responds non-monotonically to GGF intervention. In the short run it weakly declines or stays flat when the entry-expansion condition in Proposition 3(i) holds (the exploration phase). In the medium run it rises, once GGF-backed successes reveal d^* through the outcome-inference diffusion rule (the convergence phase).*

Prediction 3 (Sustained Entry and Innovation Upgrading)**.** *After GGF entry, new firm formation in the same industry-region rises in a sustained way, and the share of high-R&D firms grows over the medium term. Because Quality Filtering (Proposition 2) raises the potential productivity of the entrant cohort even after the GGF has exited, the gains in entry and innovation outlast the active life of the fund.*

Prediction 4 (Conditions for GGF Effectiveness)**.** *The crowd-in and entry effects of GGF are increasing in (i) the geographic scope, the local entrepreneurial pool size and talent quality, and decreasing in (ii) frictions that reduce the GGF's effective enhancement capability $\bar{b}$ or raise its*

effective screening cost c. In environments where these conditions are not met, GGF intervention may fail to crowd in private capital, or even crowd out private activity through general-equilibrium wage and resource competition.

These four predictions translate the model's mechanisms into testable hypotheses, and they organize the empirical analysis that follows.

4 Empirical Evidence

4.1 Suggestive Evidence

In this section, we present suggestive evidence about the impacts of central GGFs' investments on industry-province outcomes including: future PVCs' investments, long-term new firm entries, and long-run R&D activities. Specifically, we estimate the following panel regression:

$$Y_{i,j,t+\tau} = \beta IHS(Central_{i,j,t}) + \mathbf{\Gamma'} \boldsymbol{X}_{i,j,t} + \alpha_t + \theta_i + \gamma_j + \varepsilon_{i,j,t} \tag{2}$$

where $Y_{i,j,t+\tau}$ is the outcome variable for industry i located in province j in year $t+\tau$ (τ years after the fund establishment year t); $IHS(Central_{i,j,t})$ is the inverse hyperbolic sine transformation of the number of firms in industry i located in province j that are funded by central GGFs established in year t; $\boldsymbol{X}_{i,j,t}$ is a set of control variables including: province j's ln(GDP), average wage, population, and number of universities in year t and average R&D intensity, capital intensity, profit margin, and sales growth of firms located in province j in industry i (the detailed definitions of these variables are presented in Table A1 in Appendix); θ_i is the industry fixed effects; γ_j is the province fixed effects; and α_t is the fund establishment year fixed effects. Robust standard errors are double-clustered at industry and province level.

4.1.1 Private capital

In Panel A of Table 3, we examine the impact of central GGF investments on subsequent PVC investments. Specifically, at the industry-province level, we investigate how the number of firms

funded by GGFs is associated with the number of firms funded exclusively by PVCs established one to three years later than the GGFs. Because our dependent variable, $Private_{i,j,t}$, is count data with a high frequency of zeros, we follow Cohn, Liu, and Wardlaw (2022) and estimate both Poisson regressions and OLS regressions using the inverse hyperbolic sine transformation of the dependent variable, $IHS(Private_{i,j,t})$. Columns (1) and (2) report OLS estimates where the dependent variable is the inverse hyperbolic sine transformation of the number of PVC-funded firms, measured at $t+1$ and $t+3$, respectively. Columns (3) and (4) report corresponding Poisson estimates using the raw count outcome.

Across all specifications, we find strong and consistent evidence that central GGF investments crowd in subsequent private capital. In columns (1) - (2), the coefficients on *IHS(Central)* are positive and highly significant, indicating that industry-province cells receiving any central GGF investment experience a substantial increase in the number of PVC-funded startups in following years. The results are quantitatively similar in the Poisson specifications in columns (3) - (4), alleviating concerns that the findings are driven by functional form assumptions or the treatment of zeros. To translate the coefficient into an economic magnitude, we consider a case in which central GGF investment moves from the sample mean of 2.41 firms per industry-province (Table 1) to 3.41 firms—a roughly 40% increase from the mean.[8] The implied effect is a 20% increase in expected subsequent PVC investment at $t+1$ and a 14% increase at $t+3$.

Taken together, Panel A provides robust evidence that central GGF investments are associated with economically and statistically significant increases in subsequent private venture capital activity, consistent with a crowd-in rather than a crowd-out effect.

4.1.2 New firm entry

Panel B of Table 3 examines whether central GGF investments generate longer-run real effects on entrepreneurial activity, measured by new firm entry at the industry-province level. The dependent

[8]The implied multiplicative effect under Poisson is $\exp(\hat{\beta} \times \Delta\text{IHS}) - 1$, where $\Delta\text{IHS} = \sinh^{-1}(3.41) - \sinh^{-1}(2.41) = 0.328$. For column (3), $\exp(0.542 \times 0.328) - 1 \approx 0.195$, i.e., a 19.5% increase. We adopt this magnitude convention throughout the paper.

variable is the number of newly established firms in year $t+3$ and $t+5$, estimated via OLS on $Ln(Entry)$ (columns 1-2) and Poisson (columns 3-4). The Poisson coefficients on $IHS(Central)$ are 0.186 at $t+3$ and 0.164 at $t+5$, both significant at the 1% level. Applying the same magnitude convention as in Panel A (ΔIHS $= 0.328$), one additional central GGF investment at the sample mean is associated with a 6.3% increase in expected new firm entry at $t+3$ and a 5.5% increase at $t+5$.[9] The OLS estimates in columns (1) and (2) are quantitatively similar ($\hat{\beta} = 0.226$ and 0.242 on $Ln(Entry+1)$). Taken together, Panel B provides consistent evidence that central GGF investments are associated not only with increased private venture financing in the short run but also with a sustained increase in real entrepreneurial entry over the subsequent three to five years, consistent with central GGFs fostering a more supportive local entrepreneurial ecosystem.

4.1.3 R&D activity

Panel C of Table 3 examines whether central GGF investments are associated with longer-run innovation-related outcomes at the industry-province level. We consider two measures. The first is $Top_R\&D_{t+k}$, defined as the share of firms in an industry-province whose R&D intensity falls in the top decile of that industry, measured five and seven years after the central GGF investment. The second is $Tech_Share_{t+k}$, defined as the share of technology-oriented firms, those with positive R&D expenditures or receiving high-tech government subsidies, measured at horizons of $t+5$ and $t+7$. All specifications are estimated using OLS.

Across all columns, we find a positive and statistically significant relationship between central GGF investment and subsequent industrial upgrading. Columns (1) and (2) show that central GGFs significantly increase the concentration of high-R&D firms: the coefficients on *IHS(Central)* are 0.048 (significant at the 1% level) and 0.034 (significant at the 5% level) for the five- and seven-year horizons, respectively. Applying the same magnitude convention (ΔIHS $= 0.328$), one additional central GGF investment at the sample mean implies a 1.6 percentage point increase in the share of

[9] For column (3), $\exp(0.186 \times 0.328) - 1 \approx 0.063$; for column (4), $\exp(0.164 \times 0.328) - 1 \approx 0.055$.

top-decile R&D firms at $t+5$ and a 1.1 percentage point increase at $t+7$.[10] The coefficients for *Tech_Share* (columns 3-4) are 0.020 at both horizons (significant at the 1% level), implying a 0.7 percentage point increase.

Taken together, Panel C suggests that central GGF investments are not only associated with greater private capital inflows and higher firm entry, but also with longer-run shifts in real economic activity toward more R&D-intensive and technology-oriented firms. These patterns are consistent with central GGFs shaping the technological trajectory of local entrepreneurial ecosystems rather than merely increasing the volume of financing.

4.1.4 Robustness checks

In Table IA1 in the Internet Appendix, we examine whether our baseline results in Table 3 are robust to alternative measures of GGF investments and different fixed effects. In Panel A of Table IA1, we measure the central GGFs' investments as the number of firms funded solely by central GGFs. This measure excludes firms that are invested by both central and local GGFs. As shown in Panel A, the results remain economically and statistically significant when restricting the sample to these solely central-funded investments. The coefficients on *IHS(Central_Solely)* are positive and significant at the 1% level for private capital (columns 12), new firm entry (columns 34), and innovation (columns 58). This confirms that our findings are driven by the direct impact of central GGFs rather than by co-investments with local government funds.

In Panel B of Table IA1, we include alternative fixed effects instead of the industry, province, and year fixed effects used in our baseline specifications. In columns (1), (3), (5), and (7), the fixed effects are province fixed effects and year $\times$ industry fixed effects. In columns (2), (4), (6), and (8), the fixed effects are industry fixed effects and year $\times$ province fixed effects. The baseline results continue to hold with these alternative fixed effects. This provides reassurance that our results are not driven by unobserved time-varying industry shocks (such as sector-specific booms) or unobserved time-varying local shocks (such as regional policy changes or economic cycles).

[10] For Panel C, the implied change in the share outcome is $\hat{\beta} \times \Delta$IHS in level units (not exponentiated), since the dependent variable is itself a share rather than a count.

4.2 Instrumental Variable for Central GGF: Shift-Share Design

To further mitigate concerns that unobserved industry-province characteristics may attract both GGF and PVC fundings and draw more entrepreneurs into a given industry-province, we construct an instrumental variable based on a shift-share design that is specifically for central GGF measures. For the pre-determined share, for each industry, we compute the share of sales accounted for by each province in 2007, the earliest year in our tax survey data that allows us to measure each province's contribution to each industry. Specially, we define the share of sales in 2007 for the province j in the industry i as follows:

$$Share_{i,j,2007} = \frac{Sales_{i,j,2007}}{Sales_{i,2007}}, \tag{3}$$

where $Sales_{i,j,2007}$ is the total sales of firms that were in the 2007 tax survey sample in industry i and is located in province j; and $Sales_{i,2007}$ represents the total sales of firms that were in the 2007 tax survey sample in industry i.

We then interact this pre-determined provincial share with the annual total amount of capital raised by GGFs that target specific industries. Specially, we define the following shift-share instrument:

$$Shift_Share_{i,j,t} = Raised_{i,t} \times Share_{i,j,2007} \tag{4}$$

where $Raised_{i,t}$ in the total amount of funding raised in year t by GGFs targeting industry i.

It is worth noting that our instrument, unlike standard shift-share designs in which each observational unit is exposed to multiple shifts in a given period, is a special case with only a single shift per unit-period. This structure is similar to the one-shock Bartik specification in Nunn and Qian (2014). In addition, because we only have one single exposure share and this share, by definition, smaller than one (i.e., incomplete share), we follow the guidance in Borusyak, Hull, and Jaravel (2025) and include interactions between the shares and period indicators in our Two-Stage Least Squares (2SLS) specification to control for potential time-varying effects of the shares. Specifi-

cally, we estimate the following 2SLS specification for the sample period of 2008 - 2016:

$$IHS(Central_{i,j,t}) = \beta_1 Shift_Share_{i,j,t} + \sum_{t=2008}^{2016} \tau_t(\gamma_t \times Share_{i,j,2007}) + \mathbf{\Gamma'} \boldsymbol{X}_{i,j,t} + \alpha_j + \lambda_i + \gamma_t + \varepsilon_{i,j,t} \tag{5}$$

$$\tilde{Y}_{i,j,t} = \beta_2 \widehat{IHS(Central_{i,j,t})} + \sum_{t=2008}^{2016} \tau_t(\gamma_t \times Share_{i,j,2007}) + \mathbf{\Gamma'} \boldsymbol{X}_{i,j,t} + \alpha_j + \lambda_i + \gamma_t + \varepsilon_{i,j,t} \tag{6}$$

where $Shift_Share_{i,j,t}$ is the excluded instrument. $\tilde{Y}_{i,j,t}$ represents the outcome variable for industry i in province j at time t. $\widehat{IHS(Central_{i,j,t})}$ denotes the fitted values from the first stage. $Share_{i,j,2007}$ represents the baseline industry shares in 2007. By interacting these with year fixed effects (γ_t), we control for time-varying effects of the initial exposure, ensuring identification stems from the exogenous shocks. $\boldsymbol{X}_{i,j,t}$ is a vector of time-varying control variables with coefficient vector $\mathbf{\Gamma}$, α_j, λ_i, and γ_t denote province, industry, and year fixed effects, respectively. $\varepsilon_{i,j,t}$ is the idiosyncratic error term. The robust standard errors are clustered at industry×province level.

4.3 Identification

Our identification relies on the idea that variation in central GGF fundraising over time reflects shifts in nationwide policy priorities, rather than responses to localized investment conditions. This identification logic is analogous to standard Bartik (shiftshare) designs, in which national-level shocks are interacted with predetermined local exposure shares, and identification hinges on the exogeneity of the aggregate shocks rather than the exposure shares themselves.[11] In particular, we assume that central GGFs do not systematically raise more funds in years when industries that are heavily concentrated in certain provinces are about to experience unusually strong growth in private investment or entrepreneurial activity.

This assumption is natural in our setting. Central GGFs are established by national ministries

[11] See Bartik (1991) for the original formulation of the shiftshare design. For recent discussions of identification assumptions and interpretation, see Goldsmith-Pinkham, Sorkin, and Swift (2020) and Borusyak, Hull, and Jaravel (2022).

and centrally controlled SOEs to implement broad, national industrial strategies, and their fundraising decisions are made at the central level, largely independent of provinceindustryspecific shocks. It is therefore unlikely that unobserved local investment opportunities or technological improvements in a particular provinceindustry pair would trigger coordinated increases in fundraising by multiple central government entities targeting that specific location and industry. To examine this, we look into Strategic Emerging Industries Initiative.

4.3.1 Validation of identification assumption: Strategic Emerging Industries Initiative

To provide direct institutional support for this interpretation, we examine the Strategic Emerging Industries (SEI) Initiative, a flagship national industrial policy that illustrates how sectoral priorities are set at the national level and shaped by strategic considerations rather than by contemporaneous local investment conditions.

In 2012, the Chinese government officially designated seven sectors as Strategic Emerging Industries (SEIs): energy-efficient and environmental technologies, next-generation information technology, biotechnology, high-end equipment manufacturing, new energy, new materials, and new-energy vehicles. The policy document issued by the State Council articulates three primary rationales for prioritizing these industries. First, SEIs are characterized by large positive externalities. The document emphasizes that the development of these industries plays a foundational role in guiding the long-term trajectory of economic and social development.[12] Consistent with this rationale, recent studies document that industrial policy tends to target sectors with high social spillovers (e.g., Liu, 2019; Ju et al., 2024). Second, SEIs are framed as central to national security and technological sovereignty. The policy statement repeatedly stresses the importance of "independent innovation," explicitly noting that many of these industries lag behind advanced economies in their domestic innovation capabilities. Third, the initiative explicitly justifies government intervention on the grounds of private underinvestment. The statement notes that existing investment and financing mechanisms and market institutions were insufficient to support the de-

[12]The full statement is available at https://www.gov.cn/zwgk/2012-07/20/content_2187770.htm.

velopment needs of these sectors, highlighting a perceived gap between private incentives and social returns.

Together, these rationales underscore that SEI designation reflects national strategic priorities, rather than contemporaneous market performance or local growth prospects. This institutional context supports our identifying assumption that central GGF fundraising responds primarily to top-down policy shifts rather than bottom-up local conditions.

We next examine whether the empirical characteristics of SEIs are consistent with the policy rationales articulated in the initiative. Panel A of Table IA2 in the Internet Appendix compares SEIs and non-SEIs along several industry-level dimensions measured prior to the announcement of the initiative, including R&D intensity, capital intensity, market-to-book ratios, employment, and return on assets.

SEIs exhibit substantially higher R&D intensity: average R&D expenditures scaled by assets are 0.2% in SEIs, approximately twice the level observed in non-SEIs, and the difference is statistically significant at the 1% level. This pattern aligns with the policy rationale that SEIs are sectors characterized by high social returns to innovation. Indeed, the literature estimates that the average social return to R&D, reflecting externalities not internalized by private investors, is on the order of 30% (e.g., Jones and Williams, 1998).

We next examine whether SEIs are more closely tied to national security considerations. To do so, we study whether firms operating in SEIs are more likely to face sanctions by the U.S. government, which provides revealed-preference evidence of strategic importance. Panel B of Table IA2 reports the probability of appearing on U.S. sanctions lists for firms founded no later than 2016.

In the Chinese tax survey sample (Sample I), firms in non-SEIs face a sanction probability of 0.007%, whereas firms in SEIs face a probability of 0.047%, a nearly sixfold increase. In the VC/PE-backed sample (Sample II), the corresponding probabilities are 0.094% for non-SEIs and 0.191% for SEIs, implying a doubling of sanction risk. These patterns are consistent with SEIs being disproportionately concentrated in sectors that are geopolitically sensitive and strategically

important.

Panels C and D extend this analysis to firms with upstream or downstream linkages to SEIs. Firms whose suppliers or customers operate in SEIs are also significantly more likely to appear on U.S. sanctions lists. Relative to firms whose upstream or downstream industries are not SEIs, the probability of sanction increases by at least 110% (upstream exposure) and 101% (downstream exposure). These findings reinforce the interpretation that SEIs and the broader ecosystems connected to them are strategically important and disproportionately exposed to geopolitical risk.

Finally, we examine whether SEIs possessed technological leadership at the time of their designation. Following Liu and Ma (2023), we define a "lead industry" as one whose patent self-citation ratio (the share of backward citations to domestic patents) falls in the top tercile across industries. Table IA3 shows that at the time of the SEI announcement and in the two subsequent years (20132014), none of the SEIs qualify as technological leaders. Only after the launch of the initiative does the share of SEIs classified as lead industries begin to rise, reaching 25% by 20152016. In contrast, the fraction of lead industries among non-SEIs declines monotonically from 57% in 2013 to 36% in 2016.

Taken together, SEIs are characterized ex-ante by high R&D intensity, elevated geopolitical sensitivity, and technological disadvantage relative to foreign competitors. These are precisely the types of industries in which private capital is least likely to invest aggressively based solely on expected financial returns. This evidence reinforces the interpretation that central GGF fundraising reflects top-down strategic priorities rather than bottom-up market conditions, lending further credibility to our identification strategy.

4.4 The 2SLS results

In the second stage of our 2SLS analysis, we examine a suite of outcome variables to capture the multi-dimensional impact of GGF investment. Specifically, we investigate: (i) the number of firms receiving exclusive PVC investment at horizons $t+1$ and $t+3$; (ii) the scale of new firm entry at $t+3$ and $t+5$; and (iii) the province-industry share of high-innovation firms, defined as the proportion

of firms within a specific cell whose R&D intensity ranks in the top decile ($Top_R\&D_{t+\tau}$) of their respective industry at $t+5$ and $t+7$. All horizons are measured relative to the GGF establishment year t.

The 2SLS estimates are reported in Panel A of Table 4. These results closely mirror the OLS and Poisson findings in Table 3, providing additional support for our main hypothesis. In column (1), the coefficient on $\widehat{IHS(Central)}$ is 0.683 ($p < 0.01$) for $IHS(Private_{t+1})$, and remains positive and significant at the three-year horizon in column (2). These magnitudes are remarkably close to the corresponding OLS estimates in Table 3, suggesting that endogeneity in the OLS specifications is unlikely to be driving the results. For example, the estimated coefficient for the $t+1$ horizon is 0.683 (Column 1), compared with 0.649 in the corresponding OLS specification (Panel A, Column (1) of Table 3).

We also find evidence of longer-run real effects. Columns (3) and (4) show that central GGF investment significantly increases new firm entry at the three- and five-year horizon. Columns (5) and (6) further shows a positive and significant effect on the long-term innovation outcomes for $t+5$ horizon.

To gauge economic magnitudes, we apply the same convention as in Section 4 (Table 3): a one-firm increase in central GGF investment evaluated at the sample mean, corresponding to $\Delta\text{IHS}(Central) = \sinh^{-1}(3.41) - \sinh^{-1}(2.41) = 0.328$. Because the private-capital outcome is itself IHS-transformed, we evaluate the implied effect at the sample mean of $Private$ (2.73 firms, Table 1): the 2SLS estimates imply an increase from this mean to 3.46 firms at $t+1$ and to 3.13 firms at $t+3$, that is, a 27% and a 14% increase relative to the sample mean of private VC investment, respectively.[13] For new firm entry, where the outcome is in logs, the estimates imply a 16% increase at $t+3$ and a 21% increase at $t+5$ ($\hat{\beta} \times \Delta\text{IHS}$).[14] For the share of top-decile R&D firms, whose outcome is a level share, the implied effect is $\hat{\beta} \times \Delta\text{IHS}$: an 8.0 percentage point in-

[13] For an IHS-transformed outcome with coefficient $\hat{\beta}$, the implied level is $\sinh\big(\sinh^{-1}(\overline{Private}) + \hat{\beta} \times \Delta\text{IHS}\big)$, with $\overline{Private} = 2.73$, $\sinh^{-1}(x) = \ln(x + \sqrt{x^2+1})$, and $\Delta\text{IHS} = 0.328$. Column (1): $\sinh(\sinh^{-1}(2.73) + 0.683 \times 0.328) = 3.46$, a 26.6% increase; column (2): 0.390 implies 3.13 firms, a 14.5% increase.

[14] $0.503 \times 0.328 \approx 0.165$ and $0.648 \times 0.328 \approx 0.213$.

crease at $t+5$.[15] These IV magnitudes are comparable to, though somewhat larger than, the OLS and Poisson estimates in Table 3, consistent with the larger 2SLS coefficients and a local-average-treatment-effect interpretation.

While the results remain robust across most horizons, the estimates for the $t+7$ horizon are less precise, likely due to the inherent reduction in sample size as the panel extends. The first-stage results in Panel B show that instrument strength deteriorates, where the F-statistics fall below conventional thresholds.

Overall, these results establish a causal link between central GGF investment and subsequent increase in both the volume of private investment and the entry of new firms in the target province-industry cell, alongside a higher concentration of R&D-intensive firms.

5 Mechanism

In previous sections, we show that controlling for government subsidies and tax breaks, GGF investments crowd in private capital investments to the GGF-invested province-industries. There are many possible reasons behind these findings. As discussed in Section 2.2, GGF investments can serve as a pre-commitment device and send strong signals to investors about governments' future support to the targeted industries. In this section, we explore another possible mechanism that is potentially more important for emerging industries with high technological uncertainty.

GGFs tend to favor emerging industries with intensive R&D investments as evidenced in Table IA2. However, due to the high uncertainty surrounding which technological paths will prove viable, a significant portion of these R&D investments may generate no returns. As a result, private capital is generally hesitant to enter these industries until the technological uncertainty decreases. If early-stage GGF investments in these sectors can help reduce this uncertainty, private capital will naturally be more inclined to follow.

In this section, we examine whether GGF investments indeed reduce the technological un-

[15] $0.245 \times 0.328 \approx 0.080$.

certainty of the invested industries. We argue that the technological uncertainty is high if the technological paths pursued by firms within the same industry are more diverse, and vice versa.

5.1 Measure patent similarity

We thus follow Bloom, Schankerman, and Van Reenen (2013) and use the concept of cosine similarity, denoted as *Patent_Similarity*, to measure technological uncertainty of a given industry. Bloom, Schankerman, and Van Reenen (2013) use this concept to gauge the distance of two firms in the technology space. Specifically, *Patent_Similarity* is defined in terms of the share of a firm's patents in each IPC class for all the firm pairs in a given industry and a given year in which patents are filed. The position of a firm in the technology space is defined by a $N \times 1$ vector $T_i \equiv [T_{i,\tau}]_{\tau=1,\cdots,N}$, where N is the total number of IPC classes, $T_{i,\tau}$ is the share of patents of firm i in technology class τ. The technology closeness between firm i and j, $i \neq j$, is measured by the uncentered correlation between T_i and T_j (Jaffe, 1986) as

$$d_{ij} = \frac{T_i' T_j}{\left(T_i' T_i\right)^{1/2} \left(T_j' T_j\right)^{1/2}}. \tag{7}$$

A higher d_{ij} indicates higher similarity between the technological paths adopted by firms i and j. We then define the similarity of patents filed by firms in industry g as the average similarity of each pair of firms in industry g,

$$Patent_Similarity_g = \sum_i = 1_g^M d_{ij}/M_g, \tag{8}$$

where M_g is number of firms in industry g. $Patent_Similarity_g$ is higher if patents of these firms are more similar. We expect that *Patent_Ratio* becomes smaller as the technological uncertainty of an industry is reduced.

As a piece of anecdotal evidence, the Chinese government began formally supporting the electric vehicle (EV) industry in 2009 through its "New Energy Vehicle" (NEV) program. This policy

initiative catalyzed a wave of innovation and investment in the EV sector, particularly within the automobile industry.[16]

Figure 4 illustrates this trend by showing the evolution of three key metrics over time for startups in the automobile industry. The top panel plots (1) the share of EV-related patents and (2) the cosine similarity of patents, a measure of technological relatedness that captures how closely a firm's patent portfolio aligns with the broader EV space. Both metrics show a rising trend over the 2009-2018 period, suggesting increasing specialization and convergence toward EV technologies. The bottom panel shows the number of automobile startups receiving investment from GGFs. Investment sharply increased after 2014, coinciding with the growing technological focus on EVs. These patterns together suggest how government investment may help reduce technological uncertainty and guide private capital toward emerging innovations.

5.2 GGF vs. Patent Similarity

Table 5 reports OLS (Panel A) and 2SLS (Panel B) estimates examining how central GGF investment affects industry-level patent similarity at horizons from one to five years ahead. Panel A presents OLS estimates. The coefficients on $IHS(Central)$ are positive and statistically significant at horizons $t+2$ through $t+4$, indicating that GGF investment is associated with an increase in patent similarity in the medium run. Panel B presents 2SLS estimates using the same shiftshare instrument as in Table 4. The IV results reveal a more informative dynamic pattern. At short horizons ($t+1$ and $t+2$), the coefficients are negative or statistically insignificant. Beginning at the three-year horizon, the estimates become positive and highly significant. The magnitudes increase at $t+3$ through $t+5$, indicating a subsequent convergence of innovation activity toward more similar technological trajectories.

This dynamic pattern is economically intuitive. Early-stage GGF investment appears to encourage experimentation across a broader set of technological paths. Over time, as learning and

[16] For instance, see https://www.technologyreview.com/2023/02/21/1068880/how-did-china-dominate-electric-cars-policy/

selection occur, unsuccessful trajectories are abandoned and firms innovation strategies converge, leading to higher patent similarity and lower technological uncertainty. This mechanism is consistent with theories of innovation under uncertainty in which public funding promotes exploratory search and accelerates information production, thereby facilitating private-sector entry once uncertainty is resolved.

Taken together, the evidence in Table 5 supports the interpretation that GGF investments affect not only the quantity of investment and entry, but also the informational environment of innovative industries. By accelerating experimentation and subsequent convergence in technological direction, GGFs appear to reduce uncertainty in emerging sectors, thereby providing a plausible mechanism through which public capital crowds in private investment.

6 Policy Implementations: Local GGF Effectiveness

Our analysis so far focuses on central GGFs, for which the shift-share design provides clear identification. China's GGF system, however, also includes local GGFs established by provincial and municipal governments to implement industrial policies. Relative to central GGFs, local governments may have finer information about firms in their jurisdictions, potentially improving screening. At the same time, local GGFs are restricted to investing in local firms, which can shrink the pool of high-quality candidates. As the model in Section 3 implies, this constraint weakens the screening channel and, in turn, limits the extent to which GGFs reduce technological uncertainty in targeted industries.

In this section, we compare the impacts of central and local GGFs on industry-province outcomes including private capital, new firm entry, technological uncertainty proxied by patent similarity, and long-run innovation. We then examine how the effectiveness of local GGFs varies with local economic endowments and government distortions.

6.1 Comparison of central and local GGFs

We begin by comparing the average effects of central and local GGFs on industry-province outcomes. Table 6 reports OLS estimates of the impact of GGFs on four key measures: private capital fundraising, new firm entry, patent similarity, and the share of high-tech firms. The results reveal a stark contrast between the two types of funds. Central GGFs (*IHS(Central)*) exhibit a robust and positive impact across all dimensions. As shown in columns (1) through (8), the coefficients for central GGFs are statistically significant at the 1% or 10% level for all outcomes. This suggests that central funds not only successfully crowd in private capital and new entrepreneurs but also foster innovation and reduce the technological uncertainty.

In contrast, the effectiveness of local GGFs (*IHS(Local)*) appears to be narrower in scope and less persistent. While local GGFs significantly increase private capital (columns 1-2) and stimulate new firm entry in the short run (column 3), these effects do not translate into long-term ecosystem development. Notably, the effect on new firm entry becomes statistically insignificant by year $t+5$ (column 4), whereas the effect of Central GGFs remains significant and economically large (0.195^{***}). This suggests that while local GGFs can generate an initial burst of entrepreneurial activity, they fail to sustain this momentum over the medium term. Furthermore, local GGFs do not generate significant technological spillovers: the coefficients for *Patent_Similarity* (columns 56) and *Tech_Share* (columns 7-8) are statistically insignificant. This implies that although local governments can mobilize financial resources, their geographic investment restrictions or lack of expertise may prevent them from directing capital toward the most technologically innovative firms.

Consistent with our findings on capital and entry, Guo and Zhao (2025) also document that local GGFs are associated with increased private investment and firm entry. We note, however, that establishing a causal average effect in this setting is challenging given the endogenous nature of local industrial policy. Local governments' choice of targeted industries is likely shaped by unobserved local economic conditions that correlate with PVC investment and entrepreneurial entry. Thus, we emphasize that the average effects reported here are suggestive, motivating our

subsequent focus on the heterogeneity of local GGF effectiveness.

6.2 Local endowment and government distortions

As suggested by the industrial agglomeration literature (e.g., Ellison, Glaeser, and Kerr, 2010), local endowmentssuch as infrastructure, the business ecosystem, and the talent poolcan shape the crowding-in effect of local GGFs. Similarly, government distortions like rent-seeking behaviors and bureaucratic inefficiency may undermine GGF effectiveness (e.g., Acemoglu and Verdier, 2000). In regions with limited endowments or high government distortions, it may be inefficient to attempt to attract PVCs and new ventures solely by establishing restricted local funds.

We proxy for local endowments using province-level annual GDP, the number of universities, and average wages (Figure IA1 presents the tercile rankings of provinces on these measures). To measure government distortions, we use city-level government corruption and efficiency indices from Nie et al. (2019), aggregated to the province level (Figure IA2 maps the provincial terciles for these indices). Higher index values indicate *lower* corruption and *higher* efficiency.

Table 7 examines how these regional characteristics moderate the effect of local GGFs on private capital (columns 15) and new firm entry (columns 610). The results indicate that the crowding-in effects of local GGFs are highly dependent on favorable local conditions. Columns (1) through (3) and (6) through (8) reveal that local GGFs effectively stimulate private capital and firm entry primarily in provinces with strong endowmentsspecifically, those in the middle and top terciles for GDP, university count, and average wages. Conversely, in provinces with the weakest endowments (the baseline groups), local GGFs fail to attract private capital and, in the case of low GDP and university counts, significantly crowd out new firm entry (-0.130^{**} and -0.190^{***}, respectively).

A similar pattern emerges regarding government distortions. Columns (4) and (9) show that local GGFs only successfully crowd in private capital (0.382^{***}) and entrepreneurs (0.186^{***}) in the least corrupted provinces (*Low_Corruption*). Likewise, columns (5) and (10) demonstrate that positive effects on private capital (0.464^{***}) and entry (0.275^{***}) are concentrated in the most efficient provinces (*High_Efficiency*). In the most corrupted and least efficient baseline regions,

local GGFs exhibit no ability to attract private capital and can negatively impact firm entry.

Taken together, these findings suggest that GGF is not a one-size-fits-all solution. The effectiveness of local GGFs depends critically on local endowments and government quality. In regions with limited endowments and high distortions, local GGFs fail to crowd in private capital and may even crowd out entrepreneurs.

7 Conclusion

Based on micro-level enterprise data, our paper provides a comprehensive study on the macroeconomic impacts of government guidance funds at the industry-province level and how these impacts vary with local economic endowments and government distortions. We show that GGFs established by the central government crowd in private funds in the same industry-province and the upstream and downstream of the invested industries. The industry-provinces invested by the central GGFs witness increased new firm entries even after the close of the funds, indicating long lasting effects. On the contrary, GGFs established by local governments crowd in private funds only in the same industry-provinces in general.

Utilizing China's unique two-tiered GGF system, we show that corruption level is the most critical factor that determines the efficacy of GGFs. Other factors such as local governments' efficiency level, investment opportunities measured by provincial GDP level, lower productivity measured by the number of universities in the region are also important. A key policy implication from our study is that industry policy needs to be implemented with careful matching of target industries with local economic endowments, and the (non)corruption and efficiency levels need to be low enough so that government distortion does not outweigh the benefits.

References

Acemoglu, Daron and Thierry Verdier (2000). "The choice between market failures and corruption". *American economic review* 91.1, pp. 194–211.

Aghion, Philippe, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros (2015). "Industrial Policy and Competition". *American Economic Journal: Macroeconomics* 7.4, pp. 1–32.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2025). "A practical guide to shift-share instruments". *Journal of Economic Perspectives* 39.1, pp. 181–204.

Chen, Zhao, Xian Jiang, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu (2023). "Tax policy and lumpy investment behaviour: evidence from China's VAT reform". *The Review of Economic Studies* 90.2, pp. 634–674.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu (2021). "Notching R&D investment with corporate income tax cuts in China". *American Economic Review* 111.7, pp. 2065–2100.

Cohn, Jonathan B, Zack Liu, and Malcolm I Wardlaw (2022). "Count (and count-like) data in finance". *Journal of Financial Economics* 146.2, pp. 529–551.

Colonnelli, Emanuele, Bo Li, and Ernest Liu (2024). "Investing with the government: A field experiment in China". *Journal of Political Economy* 132.1, pp. 248–294.

Ellison, Glenn, Edward L Glaeser, and William R Kerr (2010). "What causes industry agglomeration? Evidence from coagglomeration patterns". *American Economic Review* 100.3, pp. 1195–1213.

Griliches, Zvi (1991). "The search for R&D spillovers". *National Bureau of Economic Research Working Paper Series* w3768.

Guo, Jingyuan and Da Zhao (Nov. 2025). "Government Venture Capital and Local High-Tech Industries: Evidence from China". Working paper.

Hall, Bronwyn H (2002). "The financing of research and development". *Oxford review of economic policy* 18.1, pp. 35–51.

Hirschman, Albert O. (1958). *The Strategy of Economic Development*. Yale University Press.

Itskhoki, Oleg and Benjamin Moll (2019). "Optimal development policies with financial frictions". *Econometrica* 87.1, pp. 139–173.

Jaffe, Adam (1986). "Technological Opportunity and Spillovers of R&D: Evidence from Firmsí-Patents, Profits and Market Value". *American Economic Review* 76, pp. 984–1001.

Jones, Charles I and John C Williams (1998). "Measuring the social return to R&D". *The Quarterly Journal of Economics* 113.4, pp. 1119–1135.

Ju, Jiandong, Hong Ma, Zi Wang, and Xiaodong Zhu (2024). "Trade wars and industrial policy competitions: Understanding the US-China economic conflicts". *Journal of Monetary Economics* 141, pp. 42–58.

Juhász, Réka, Nathan Lane, Emily Oehlsen, and Verónica Pérez (2023). "The Who, What, When, and How of Industrial Policy: A Text-Based Approach". *Working Paper*.

Krueger, Anne O. (1990). "Government Failures in Development". *Journal of Economic Perspectives* 4.3, pp. 9–23.

Liu, Ernest (2019). "Industrial policies in production networks". *The Quarterly Journal of Economics* 134.4, pp. 1883–1948.

Liu, Ernest and Song Ma (2023). "Innovation networks and R&D allocation". w29607.

Nie, Huihua, Donglin Han, Liang Ma, and Nandiyang Zhang (2019). "Ranking of Urban Government-Business Relations in Chinese Cities (2018)". *The National Academy of Development and Strategy at Renmin University of China*.

Nunn, Nathan and Nancy Qian (2014). "US food aid and civil conflict". *American economic review* 104.6, pp. 1630–1666.

Rosenstein-Rodan, Paul (1943). "Problems of Industrialization of Eastern and South-Eastern Europe". *The Economic Journal*.

Figure 1.

This figure plots the total number of industrial policy interventions globally by year. Data source: Juhász et al., 2023 (data update till July, 2023).

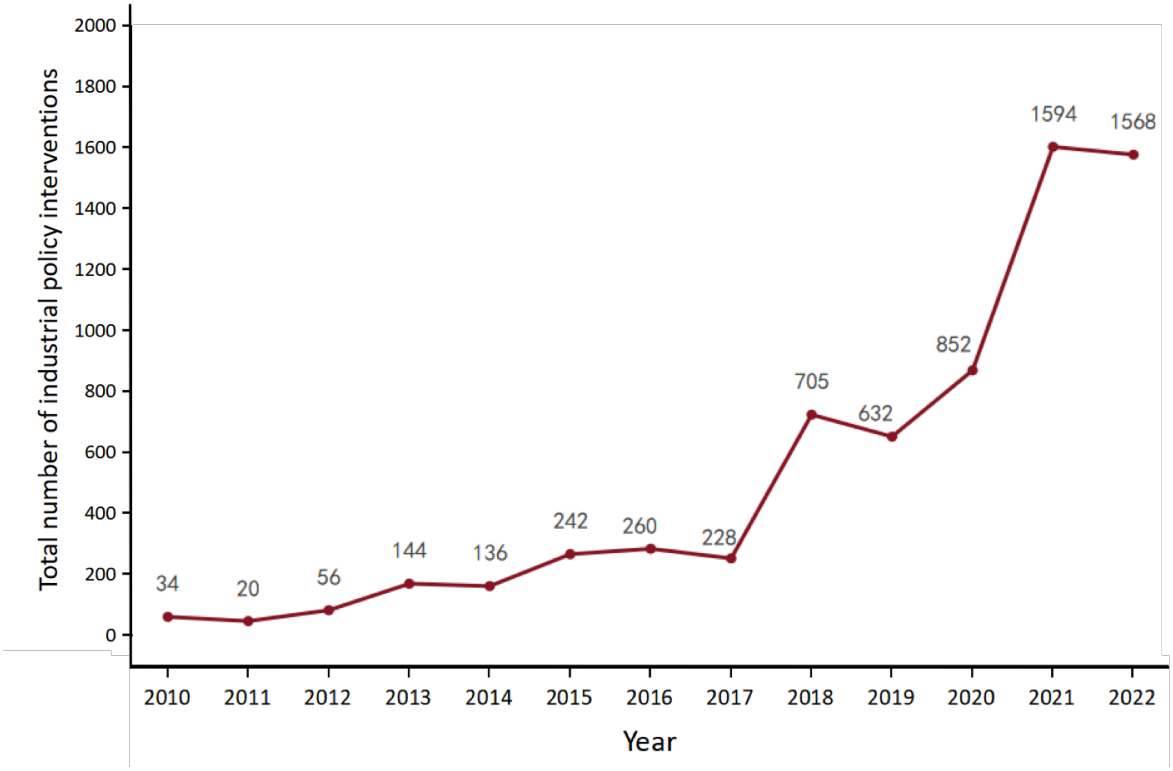


Figure 2.

This figure plots the number of total GGFs and the proportion of GGFs relative to all VC/PE funds in China by fund establishment year.

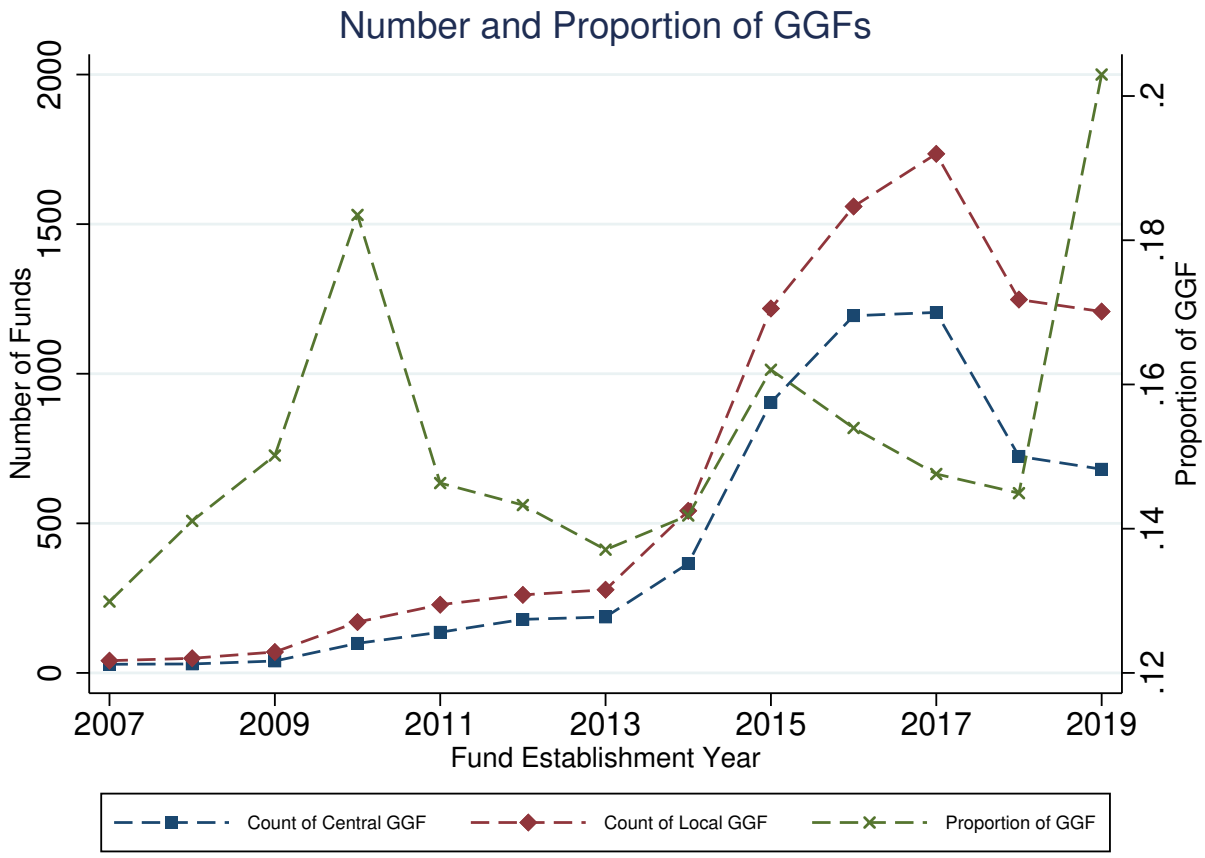


Figure 3.

This figure plots the number of firms invested by GGF and by PVC funds solely by fund establishment year.

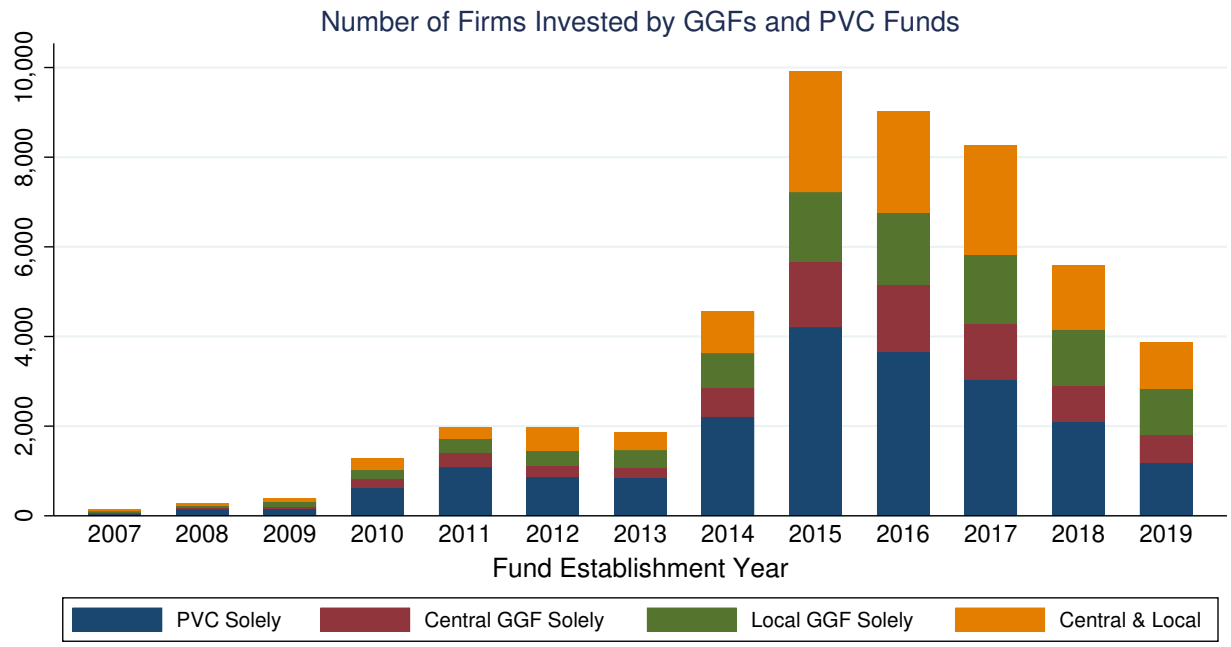
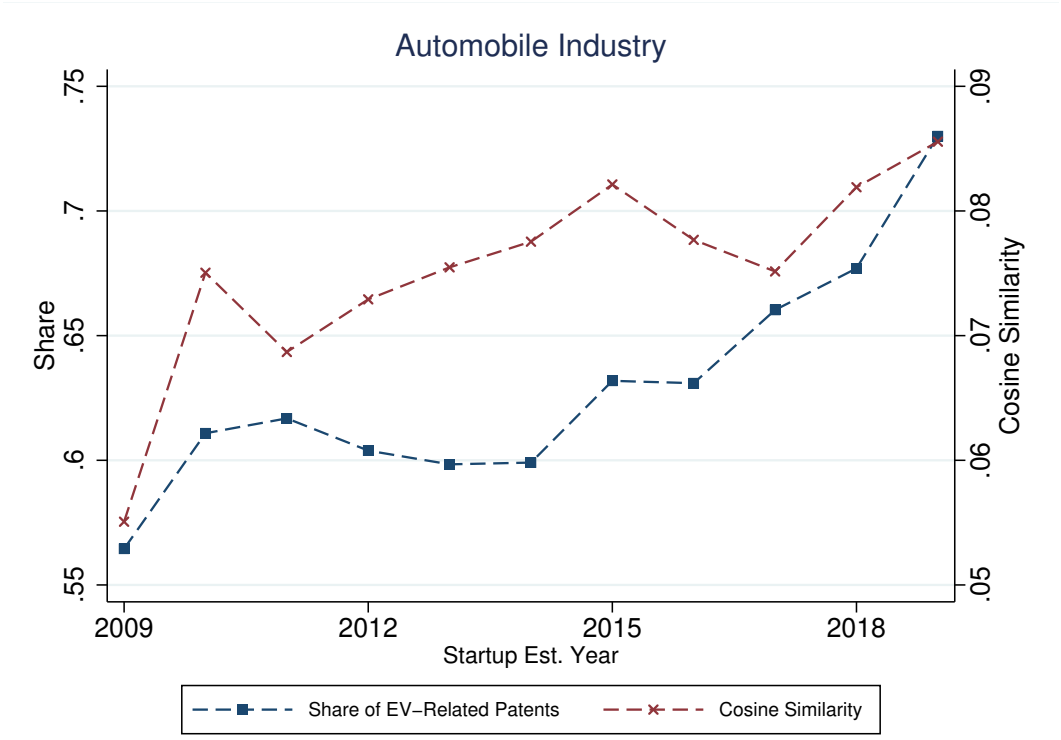


Figure 4.

The top panel of this figure plots the number of firms invested by GGF in the automobile industry and the share of patents related to electric vehicle in the same industry. The bottom panel plots the number of automobile startups receiving investment from GGFs.



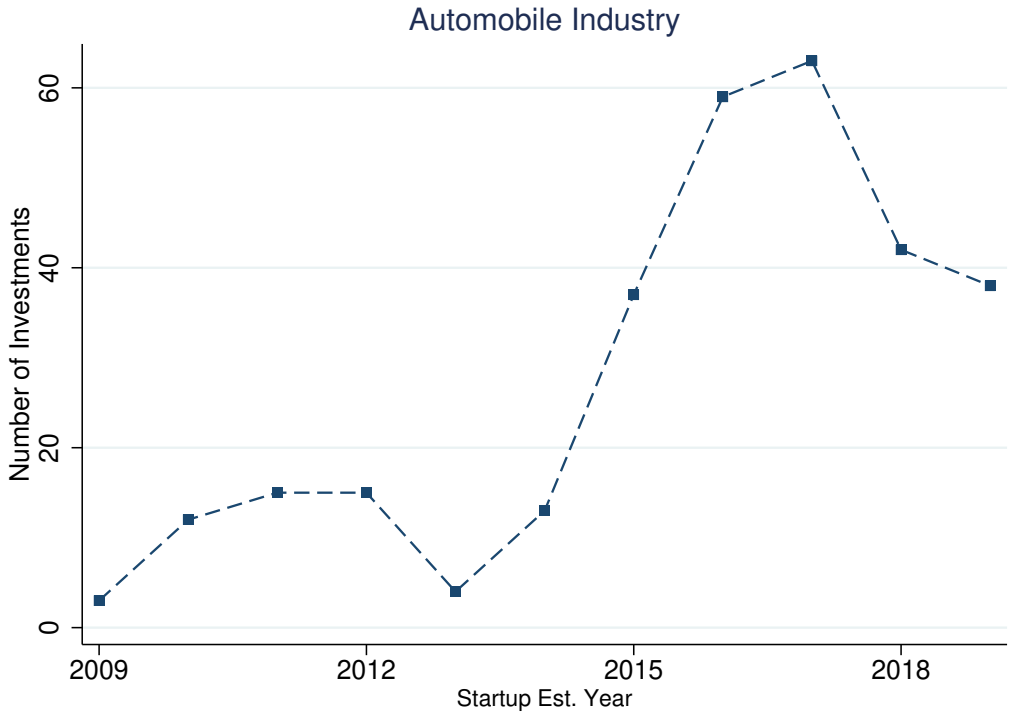


Table 1: Summary Statistics

Panel A of this table presents the summary statistics at the province-industry-year level from 2007 to 2016. Panel B presents the summary statistics at firm-year level. The details of variable definitions are in Table A1 in Appendix.

Panel A. Industry-Province Level

	Mean	Std	P1	P25	P50	P75	P99
Local	2.52	20.11	0.00	0.00	0.00	0.00	50.00
Central	2.41	19.68	0.00	0.00	0.00	1.00	43.00
Private	2.73	20.14	0.00	0.00	0.00	1.00	55.00
Entry	121.66	312.96	0.00	3.00	26.00	104.00	1500.00
R&D Intensity (%)	0.21	1.36	0.00	0.00	0.03	0.13	2.56
Capital Intensity	1.85	2.71	0.09	0.63	1.12	1.90	15.11
ROA (%)	5.13	20.43	-21.33	1.82	4.81	9.32	25.29
Market-to-Book	2.91	1.05	1.22	2.07	2.71	3.59	6.50
Subsidy (million RMB)	1.08	7.05	0.00	0.09	0.30	0.84	9.63
Age	8.39	2.68	3.50	6.63	7.98	9.70	17.36
Observations	4960						

Panel B. Firm Level

	Mean	Std	P1	P25	P50	P75	P99
Central_GGF	0.12	0.33	0.00	0.00	0.00	0.00	1.00
Local_GGF	0.13	0.33	0.00	0.00	0.00	0.00	1.00
PVC	0.14	0.34	0.00	0.00	0.00	0.00	1.00
Patent	1.85	5.85	0.00	0.00	0.00	1.00	40.00
Observations	92833						

Table 2: Subsidy vs. GGF

Panel A of this table compares key firm-level characteristics across three groups: firms receiving only government subsidies (Sub_only), firms receiving only Government Venture Capital (GGF_only), and firms receiving both (Sub_GGF). Reported variables include return on assets excluding subsidies (ROA_ExL_Subsidy), sales, total assets, average wage, R&D activity (including non-zero R&D, R&D-to-sales ratio), patent activity (including the probability of holding any patents, Non-zero Patent, and the number of patents, Num_Patent), and firm age. The final three columns present differences in means between groups, with associated significance levels. Asterisks denote statistical significance at the 1% (***), 5% (**), and 10% (*) levels. Panel B of this table reports OLS regression results estimating the effects of GGF investment and its interaction with direct government subsidies on these firm-level outcomes. All regressions include industry-year fixed effects. Standard errors are reported in parentheses and are clustered at the industry-year level. Statistical significance is denoted by *** p<0.01, ** p<0.05, and * p<0.10.

Panel A. Characteristics

	Sub_only	GGF_only	Sub_GGF	GGF_only - Sub_only	Sub_GGF - Sub_only	Sub_GGF - GGF_only
ROA_Exl_Subsidy	-0.003	0.002	0.034	0.006	0.037***	0.032***
Sales (Million)	116.252	296.524	354.421	180.272***	238.169***	57.897***
Total Asset (Million)	205.051	820.929	699.867	615.878***	494.816***	-121.062***
Average Wage (Thousands)	35.428	76.111	69.331	40.683***	33.903***	-6.780***
Non-zero R&D	0.059	0.133	0.440	0.074***	0.381***	0.307***
R&D_Sales	0.002	0.006	0.022	0.004***	0.020***	0.016***
Non-zero Patent	0.002	0.034	0.078	0.032***	0.076***	0.044***
Num_Patent	0.022	0.445	1.387	0.423***	1.365***	0.941***
Firm Age	8.638	8.402	9.498	-0.236	0.859***	1.095***

Panel B. Performances

	ROA_Exl_Sub	Sales	Asset	Wage	Non-zero R&D	R&D_Sales	Non-zero Patent	Age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GGF	-0.002	165.349***	536.049***	23.173***	0.024	0.001	0.028***	-0.688***
	(0.009)	(20.440)	(57.807)	(2.378)	(0.016)	(0.001)	(0.007)	(0.235)
GGF×Subsidy	0.034***	43.601**	-73.562	-2.271	0.269***	0.014***	0.038***	0.334
	(0.009)	(21.343)	(56.909)	(2.319)	(0.016)	(0.001)	(0.010)	(0.242)
Industry×Year FE	Y	Y	Y	Y	Y	Y	Y	Y
N	3429573	3547482	3495858	3221056	3547482	3478576	3547482	2494003
Adj. R^2	0.018	0.063	0.217	0.183	0.117	0.116	0.055	0.143

Table 3: The Impact of Central GGF Funds: Suggestive Evidence

This table presents evidence on the impact of Central GGF funds on the industry-province in which they invest. Panel A examines whether GGF investments crowd in private VC funds. The dependent variable, *IHS(Private$_{t+1}$)*, is the inverse hyperbolic sine transformation of the number of firms funded by private VCs that are established in year $t+1$. Panel B estimates long-term impacts on new firm entry, while Panel C investigates long-run R&D activities. The dependent variable, *Top_R&D*$_{t+5}$, is the proportion of firms in an industry-province with *R&D Intensity* in the top 10th percentile of that industry in year $t+5$. *Tech_Share*$_{t+5}$ is the share of technology firms (defined as firms with positive R&D expenditures or receiving high-tech government subsidies) in an industry-province in year $t+5$. The control variables include province-level ln(GDP), average wage, and number of universities and industry-level average R&D intensity, capital intensity, ROA, ln(subsidy), average age, and market-to-book value. For all specifications, robust standard errors are clustered at industry × province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively. The details of variable definitions are in Table A1 in Appendix.

Panel A. Private Capital

	OLS		Poisson	
	IHS(Private$_{t+1}$)	IHS(Private$_{t+3}$)	Private$_{t+1}$	Private$_{t+3}$
	(1)	(2)	(3)	(4)
IHS(Central)	0.649***	0.518***	0.542***	0.411***
	(0.027)	(0.031)	(0.043)	(0.044)
Ln(Subsidy)	0.007	0.013	0.111***	0.103***
	(0.006)	(0.009)	(0.028)	(0.033)
Controls	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
N	4654	4297	4654	4297
Adj. R^2	0.72	0.68		
Pseudo R^2			0.89	0.86

Panel B. New Firm Entry

	OLS		Poisson	
	Ln(Entry$_{t+3}$)	Ln(Entry$_{t+5}$)	Entry$_{t+3}$	Entry$_{t+5}$
	(1)	(2)	(3)	(4)
IHS(Central)	0.226***	0.242***	0.186***	0.164***
	(0.030)	(0.037)	(0.025)	(0.026)
Ln(Subsidy)	0.031***	0.036***	0.025	0.041**
	(0.012)	(0.013)	(0.018)	(0.016)
Controls	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
N	4599	3699	4780	3816
Adj. R^2	0.85	0.85		
Pseudo R^2			0.94	0.94

Panel C. R&D

	Top_R&D$_{t+5}$	Top_R&D$_{t+7}$	Tech_Share$_{t+5}$	Tech_Share$_{t+7}$
	(1)	(2)	(3)	(4)
IHS(Central)	0.048***	0.034**	0.020***	0.020***
	(0.007)	(0.017)	(0.005)	(0.007)
Ln(Subsidy)	0.002	0.001	0.000	0.002
	(0.003)	(0.003)	(0.002)	(0.002)
Controls	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
N	2357	1397	2357	1397
Adj. R^2	0.72	0.70	0.84	0.87

Table 4: The Impact of GGF Funds: 2SLS Estimates

This table presents 2SLS results with an instrumental variable *Shift_Share*, defined as the share of sales in 2007 from a province-industry within that industry that is interacted with the total amount of capital raised by GGF funds that target that industry. Panel A presents the second-stage results examining the effects of central GGFs on industry-province outcomes. Panel B shows the first-stage estimates. For all specifications, following Borusyak, Hull, and Jaravel (2025), control variables also include the interaction between year indicators and the shares used to construct the IV. Robust standard errors are clustered at Industry×Province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively.

Panel A. 2SLS: Second Stage

	IHS(Private$_{t+1}$)	IHS(Private$_{t+3}$)	Ln(Entry$_{t+3}$)	Ln(Entry$_{t+5}$)	Top_R&D$_{t+5}$	Tech_Share$_{t+5}$	Top_R&D$_{t+7}$	Tech_Share$_{t+7}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IHS(Central)	0.683***	0.390***	0.503***	0.648**	0.245***	0.072**	1.030*	0.077
	(0.064)	(0.094)	(0.128)	(0.273)	(0.050)	(0.034)	(0.605)	(0.150)
Ln(Subsidy)	0.004	0.016	0.023*	0.022	-0.002	0.002	0.012	0.001
	(0.007)	(0.010)	(0.012)	(0.014)	(0.004)	(0.002)	(0.011)	(0.003)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y
N	4219	3862	4190	3288	1922	1922	962	962
F-Stats	39.9	29.2	45.9	9.5	25.8	25.8	5.6	5.6

Panel B. 2SLS: First Stage

	IHS(Central)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shift_Share	0.0007***	0.0006***	0.0007***	0.0004***	0.0031***	0.0031***	0.0015**	0.0015**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Ln(Subsidy)	0.0245***	0.0243**	0.0221**	0.0194**	0.0082	0.0082	-0.0098	-0.0098
	(0.0091)	(0.0099)	(0.0091)	(0.0083)	(0.0080)	(0.0080)	(0.0075)	(0.0075)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y
N	4219	3862	4190	3288	1922	1922	962	962
F-Stats	39.9	29.2	45.9	9.5	25.8	25.8	5.6	5.6

Table 5: GGFs and Patent Similarity

This table presents OLS (Panel A) and 2SLS (Panel B) estimates examining the impacts of GGFs on industry patent similarity. The dependent variable *Patent_Similarity*, is defined as the degree of similarity among patents filed by firms within the same industry. Lower values of *Patent_Similarity* indicate greater technological uncertainty or dispersion in innovation direction. The control variables include industry-level average R&D intensity, capital intensity, ROA, ln(subsidy), average age, and market-to-book value. Robust standard errors are clustered at Industry×Province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively.

Panel A. OLS Estimates

	$\text{Patent_Similarity}_{t+1}$	$\text{Patent_Similarity}_{t+2}$	$\text{Patent_Similarity}_{t+3}$	$\text{Patent_Similarity}_{t+4}$	$\text{Patent_Similarity}_{t+5}$
	(1)	(2)	(3)	(4)	(5)
IHS(Central)	0.001	0.001*	0.003***	0.002**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ln(Subsidy)	0.002***	0.001	0.002***	0.001**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y
N	2682	2714	2713	2373	2033
Adj. R^2	0.8	0.8	0.8	0.8	0.8

Panel B. 2SLS Estimates

	$\text{Patent_Similarity}_{t+1}$	$\text{Patent_Similarity}_{t+2}$	$\text{Patent_Similarity}_{t+3}$	$\text{Patent_Similarity}_{t+4}$	$\text{Patent_Similarity}_{t+5}$
	(1)	(2)	(3)	(4)	(5)
IHS(Central)	-0.004	-0.005*	0.009***	0.044***	0.042*
	(0.005)	(0.003)	(0.004)	(0.011)	(0.023)
Ln(Subsidy)	0.002***	0.001	0.002**	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y
N	2682	2714	2713	2373	2033
F-stats	113.5	112.2	112.6	26.7	8.5

Table 6: Local GGFs

This table reports the impact of local GGF funds on four industry-province outcomes: private capital, new firm entry, R&D, and patent similarity. For all specifications, robust standard errors are clustered at industry $\times$ province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively. The details of variable definitions are in Table A1 in Appendix.

	IHS(Private)		Ln(Entry)		Patent_Similarity		Tech_Share	
	$t+1$	$t+3$	$t+3$	$t+5$	$t+3$	$t+5$	$t+5$	$t+7$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IHS(Central)	0.401***	0.277***	0.126***	0.195***	0.006***	0.001	0.022***	0.022*
	(0.030)	(0.038)	(0.041)	(0.054)	(0.002)	(0.001)	(0.008)	(0.012)
IHS(Local)	0.300***	0.288***	0.120***	0.057	-0.003*	-0.001	-0.002	-0.003
	(0.029)	(0.035)	(0.037)	(0.047)	(0.002)	(0.001)	(0.008)	(0.013)
Ln(Subsidy)	0.007	0.013	0.031***	0.036***	0.002***	0.002***	0.000	0.002
	(0.006)	(0.009)	(0.012)	(0.013)	(0.001)	(0.001)	(0.002)	(0.002)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y
N	4654	4297	4599	3699	2713	2033	2357	1397
Adj. R^2	0.738	0.690	0.848	0.849	0.798	0.848	0.839	0.868

Table 7: Local Endowment and Government Distortions

table examines how local endowments and government distortions affect the ability of local GGFs to attract private capital and entrepreneurs. Local investment opportunities are proxied by a provinces annual GDP, number of universities, and average local wage. Government distortions are measured using corruption and bureaucratic efficiency indices from Nie et al. (2019), where higher values indicate lower corruption and higher efficiency, respectively. *High* (*Mid*) is a dummy variable equal to one if a given proxy (e.g., GDP) falls in the top (middle) tercile for a province in a given year. Province-level controls include log(GDP), average wage, and number of universities. Industry-level controls include average RD intensity, capital intensity, ROA, log(subsidy), average age, and market-to-book value. Robust standard errors are clustered at industry × province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively. The details of variable definitions are in Table A1 in Appendix.

	IHS(Private$_{t+1}$)					Ln(Entry$_{t+3}$)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IHS(Central)	0.356***	0.371***	0.368***	0.348***	0.350***	0.096**	0.101**	0.118***	0.094**	0.093**
	(0.030)	(0.031)	(0.029)	(0.030)	(0.030)	(0.042)	(0.042)	(0.041)	(0.041)	(0.041)
IHS(Local)	-0.043	-0.005	0.090**	0.043	-0.041	-0.130**	-0.190***	0.031	0.014	-0.071
	(0.046)	(0.062)	(0.037)	(0.035)	(0.041)	(0.054)	(0.062)	(0.041)	(0.052)	(0.061)
Ln(Subsidy)	0.008	0.008	0.010*	0.009	0.010	0.032***	0.032***	0.032***	0.033***	0.033***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
IHS(Local)×High_GDP	0.437***					0.367***				
	(0.048)					(0.056)				
IHS(Local)×Mid_GDP	0.344***					0.153**				
	(0.059)					(0.074)				
IHS(Local)×High_Uni		0.360***					0.405***			
		(0.069)					(0.068)			
IHS(Local)×Mid_Uni		0.346***					0.268***			
		(0.075)					(0.082)			
IHS(Local)×High_Wage			0.354***					0.136***		
			(0.037)					(0.045)		
IHS(Local)×Mid_Wage			0.126***					0.081*		
			(0.043)					(0.044)		
IHS(Local)×Low_Corruption				0.382***					0.186***	
				(0.035)					(0.054)	
IHS(Local)×Mid_Corruption				0.111**					-0.020	
				(0.050)					(0.056)	
IHS(Local)×High_Efficiency					0.464***					0.275***
					(0.041)					(0.064)
IHS(Local)×Mid_Efficiency					0.229***					0.098
					(0.049)					(0.062)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	4654	4654	4654	4654	4654	4599	4599	4599	4599	4599
Adj. R^2	0.749	0.745	0.753	0.753	0.756	0.851	0.851	0.849	0.849	0.850

Internet Appendix

Appendix A Additional Data Sources

A.1 Patents

For firm quality measure, we rely on the number of granted patents. We obtain a comprehensive patent dataset from the Chinese National Intellectual Property Administration (CNIPA). The CNIPA dataset includes three types of patents filed in China: design, utility, and invention patents. Only the invention patents are comparable to the utility patents granted by the United States Patent and Trademark Office (USPTO); therefore, we focus on invention patents and refer to them simply as "patents" in this study. We merge firms from the business registration database (SAIC) the with the patents data based on a permanent legal identifier of Chinese firms, uniform social credit codes.

A.2 U.S. sanction data

We hand-collected a list of 779 Chinese firms subject to US sanction from three entity lists: the Bureau of Industry and Security (BIS) list, the Non-SDN Communist Chinese Military Companies (NS-CCMC) list effective as of June 3, 2021, and the Chinese Military-Industrial Complex (CMIC) list effective as of February 15, 2022. The sanctions on these Chinese firms were imposed between August 1, 2018, and October 17, 2023. We merged the sanction list of Chinese firms with two samples. The first one is the Chinese tax survey from 2007 to 2017. We merged 324 firms from the sanction list to the tax survey sample. The second one contains the firms invested by VC/PE. We merged 153 firms from the sanction list to the that sample.

A.3 Control variables at province-industry level

We construct control variables at province-industry level by using the National Tax Survey Database of the Chinese State Administration of Tax (SAT) from 2007 to 2016.[17] The data include firms'

[17]This database is also used in Chen et al. (2021) and Chen et al. (2023).

accounting information such as sales, profits, total assets, employment, and R&D expenses. Moreover, SAT also reports whether and how much subsidy a firms receives from the government, which is particularly relevant in our study. We then aggregate these variables at province-industry level as our control variables.

We restrict the sample to have non-negative sales and assets. The final sample contains a comprehensive list of private and publicly-listed firms with about 6.6 million firm-year observations from 2007 to 2016 and about 2.3 million distinct firms.

Since private firms do not have available market values, we use the market value of firms listed in stock exchanges in China. The market value data are extracted from the China Stock Market and Accounting Research Database (CSMAR), one of the major data providers of Chinese public firms. We then aggregate the firm-level information at industry-province level to control for a industry's development in a given province. The industry classification is based on 20 business sectors, consistent with the classification used in new firm entry data.

A.3.1 Control variables at province level

We construct control variables at province level by using the Statistical Yearbook published by National Bureau of Statistics of China (NBSC). For each year since 1994, the yearbook reports city-level statistics such as GDP, population, the number of universities, the number of employees, average local wage, etc. In our analyses, we use GDP, average local wage, and the number of universities and aggregate them at the province level as our control variables.

Appendix B Tables

Table A1: **Variable Definitions**

Variable	Definition	Source
Panel A. Industry–Province-Level Variables		
$\text{Central}_{i,j,t}$	Number of firms in industry i, province j, invested by central GGFs established in year t	SAIC
$\text{Local}_{i,j,t}$	Number of firms in industry i, province j, invested by local GGFs established in year t	SAIC
$\text{Private}_{i,j,t}$	Number of firms in industry i, province j, invested solely by PVC funds established in year t	SAIC
$\text{Entry}_{i,j,t}$	Number of newly created firms in year t, industry i, and province j	SAIC
$\text{Central_Solely}_{i,j,t}$	Firms in industry i, province j, invested by central GGFs but not by local GGFs	SAIC
$\text{Local_Solely}_{i,j,t}$	Firms in industry i, province j, invested by local GGFs but not by central GGFs	SAIC
$\text{Central_Local}_{i,j,t}$	Firms in industry i, province j, invested by both local and central GGFs	SAIC
Panel B. Industry-Level Variables		
$\text{R\&D Intensity}_{i,t}$	Average R&D expenses scaled by assets in industry i and year t	Tax Survey
$\text{Capital Intensity}_{i,t}$	Average assets scaled by employment in industry i and year t	Tax Survey
$\text{ROA}_{i,t}$	Average net income scaled by assets in industry i and year t	Tax Survey
$\text{Subsidy}_{i,t}$	Total amount of subsidy received by firms in industry i and year t	Tax Survey
$\text{Age}_{i,t}$	Average firm age in industry i and year t	Tax Survey
$\text{Market-to-Book}_{i,t}$	Average market value scaled by assets for public firms in industry i and year t	CSMAR
Panel C. Province-Level Variables		
$\text{GDP}_{j,t}$	Annual GDP in province j and year t	NBSC
$\text{Uni}_{j,t}$	Number of universities located in province j and year t	NBSC
$\text{Wage}_{j,t}$	Average wage in province j and year t	NBSC
$\text{Corruption}_{j,t}$	Corruption index for province j and year t	Nie et al. (2019)
$\text{Efficiency}_{j,t}$	Efficiency index for province j and year t	Nie et al. (2019)

Table A1: **Variable Definitions – Continued**

Variable	Definition	Source
Panel D. Firm-Level Variables		
$\text{Central_GGF}_{i,t}$	Indicator for firm i invested by central GGFs by year t	SAIC
$\text{Local_GGF}_{i,t}$	Indicator for firm i invested by local GGFs by year t	SAIC
$\text{PVC}_{i,t}$	Indicator for firm i invested by private funds by year t	SAIC
$\text{Patent}_{i,t}$	Number of granted patents applied by firm i no later than year t	SAIC

Appendix C Model Appendix

This appendix collects all technical material supporting the model in Section 3. Appendix C.1 contains the formal assumptions, closed-form expressions, and first-order conditions that the main text references but does not state in full. Appendix C.2 contains the formal policy comparison analysis. The remaining subsections contain the proofs of all lemmas, propositions, the dominance corollary, and the dynamic crowd-in prediction stated in the main text.

C.1 Technical Details of the Model

C.1.1 Technology and productivity distributions

The traditional sector uses labor under diminishing returns:

$$\tilde{y} = \tilde{l}^{\alpha}, \qquad \alpha \in (0,1),$$

with the price normalized to one. The modern sector has output $y_i = A\theta_i l_i$ for any developed good i, where $A > 0$ is a normalization. Each entrepreneur can operate at most one firm and firm size is normalized to one unit of output.

To produce, an entrepreneur must pay a fixed cost f (in labor) in period 0. Entrepreneurial ability $z \in [\underline{z}, \bar{z}]$ is privately known and shifts the productivity distribution. The distribution also depends on whether the entrepreneur pursues the viable direction:

$$\theta_i \mid (z_i, d_i = d^*) \sim F^*(x; z_i) = x^{1+bz_i}, \qquad \theta_i \mid (z_i, d_i \neq d^*) \sim F^0(x; z_i) = x^{1+\delta bz_i}, \tag{A.1}$$

where $b > 0$ and $\delta \in (0,1)$. Higher ability shifts mass toward higher productivity under either direction, and the viable direction stochastically dominates. Without information about d^*, an entrepreneur chooses each direction with equal probability $1/K$.

C.1.2 Directional research technology and indivisibility

Before entrepreneurs commit to a direction, any agent with a portfolio of multiple firms (an investor, a fund, or a planner) can invest resources $\kappa \geq 0$ in technological assessment. This investment reveals d^* with probability $q(\kappa)$, where

$$q(0) = 0, \quad q'(\kappa) > 0, \quad q''(\kappa) < 0, \quad 0 < q'(0) < \infty, \quad \lim_{\kappa \to \infty} q'(\kappa) = 0. \tag{A.2}$$

An agent who learns d^* shares this information with its own portfolio firms, who then choose $d_i = d^*$. Firms outside the agent's portfolio continue to choose directions uniformly. Equation (A.2) keeps the marginal product of the first unit of directional research finite. Because directional research is a portfolio-level fixed cost in substance (its benefit scales with the number of firms directed, while any single entrepreneur benefits through only one draw), it is uneconomical for individual entrepreneurs who control a single firm. We formalize this with the following assumptions.

Assumption 1 (Indivisibility of Directional Research)**.** *The single-firm directional premium is too small to justify any positive directional research:*

$$\sup_{z \in [\underline{z}, \bar{z}]} \left[\Pi^*(z) - \Pi^0(z) \right] \cdot \frac{K-1}{K} \cdot q'(0) \; < \; 1.$$

Assumption 2 (Portfolio Scale for Directional Research)**.** *For the PE fund, the GGF, and the social planner, the marginal portfolio value of directional research at $\kappa = 0$ exceeds its marginal cost, while the marginal value tends to zero as $\kappa \to \infty$. Hence each portfolio problem has an interior solution for κ.*

Together, Assumptions 1 and 2 say that industry-wide technological assessment, such as pilot programs, scouting, and scientific synthesis, is not worth much to a single entrepreneur but becomes valuable once its signal can be spread across a portfolio. The value $\kappa^d = 0$ in the main text should therefore be read as zero *atomistic* directional research in the decentralized economy, not

as zero private value of technological information once firms are bundled into a portfolio.

C.1.3 Period 2 information diffusion

After period 1 production, the realized productivities and the directions chosen by all entrants are publicly observable. The market then updates beliefs about d^*. We adopt the following Bayesian-inference rule:

Assumption 3 (Diffusion via Outcome Inference)**.** *At the start of period 2, each agent observes* $\{(d_i, \theta_i)\}_{i=1,\dots,m}$ *and updates beliefs about* d^* *via the maximum-likelihood rule*

$$\hat{d} \;=\; \arg \max_{d \in \{1,\dots,K\}} \; \max_{i:\, d_i = d} \theta_i .$$

New entrants and pivoting firms in period 2 then choose $d_i = \hat{d}$.

Because F^* first-order stochastically dominates F^0 for any fixed z, the maximum order statistic on the viable direction is more likely to be the global maximum. Hence $\Pr(\hat{d} = d^*)$ rises with the number and quality of period-1 entrepreneurs on d^*. Assumption 3 replaces the heuristic "competitors observe success and pivot" with an explicit inference rule. It is the channel through which directional research, once GGF-backed firms turn it into successful outcomes, diffuses to the rest of the market.

C.1.4 Closed form of the expected innovation frontier

The closed-form counterpart of Lemma 1 is the following. Conditional on the directional allocation D^* and ability profile $\{z_i\}$,

$$\mathbb{E}[\hat{\theta} \mid D^*] = 1 - \frac{1}{1 + m + b\sum_{i \in D^*} z_i + \delta b \sum_{j \notin D^*} z_j}. \tag{A.3}$$

Reallocating an entrepreneur of ability z from a wrong direction to d^* raises the frontier by

$$\frac{(1-\delta)bz}{(1+S)\big(1+S+(1-\delta)bz\big)} > 0,$$

where $S \equiv m + b\sum_{i\in D^*} z_i + \delta b \sum_{j\notin D^*} z_j$. The marginal gain is strictly increasing in z, which is the basic source of complementarity between high-ability screening and directional guidance.

C.1.5 Demand, period-2 output, and the social planner's problem

Each good has unit demand with willingness to pay p. Firms produce in period 1 if the unit cost $w_1/(A\theta_i)$ does not exceed p, earning monopoly profits for duration T. In period 2, entry becomes unrestricted, all $\{(d_i,\theta_i)\}$ are observed (Assumption 3), and competition eliminates monopoly rents.

After diffusion, only the frontier technology $\hat{\theta}$ remains active. The planner allocates labor across sectors to equalize marginal products: $\alpha(\bar{L}-l_M)^{\alpha-1} = pA\hat{\theta}$, yielding modern-sector labor $l_M(\hat{\theta})$. Aggregate output in period 2 is

$$Y_2(\hat{\theta}) = (\bar{L}-l_M)^{\alpha} + pA\hat{\theta}\cdot l_M, \tag{A.4}$$

strictly increasing in $\hat{\theta}$: $Y_2'(\hat{\theta}) = pA\cdot l_M(\hat{\theta}) > 0$. The frontier therefore enters social welfare directly, valued at the price p times the modern-sector labor it employs.

Let $\Theta(m,q)$ denote the unconditional expected frontier when m entrepreneurs enter and each has probability $q+(1-q)/K$ of pursuing d^*. By Lemma 1, Θ is strictly increasing in both m and q. The planner chooses (m,κ) to maximize

$$\max_{m,\,\kappa\geq 0} V_P(m,\kappa) = \underbrace{(\bar{L}-mf)^{\alpha}}_{\text{period 0 traditional}} + \underbrace{Y_2(\Theta(m,q(\kappa)))}_{\text{period 2 output}} - \kappa. \tag{A.5}$$

We assume V_P is jointly strictly concave, so the unique interior optimum (m^*, κ^*) satisfies

$$\text{Scale:} \quad f\alpha(\bar{L} - m^* f)^{\alpha-1} = Y_2'(\Theta^*) \cdot \Theta_m(m^*, q^*), \tag{A.6}$$

$$\text{Direction:} \quad Y_2'(\Theta^*) \cdot \Theta_q(m^*, q^*) \cdot q'(\kappa^*) = 1. \tag{A.7}$$

Equation (A.6) equates the marginal cost of entry (forgone traditional output) to its marginal social benefit through frontier improvement. Equation (A.7) equates the marginal cost of directional research to its marginal social benefit through better directional allocation. The planner uses *two* instruments, m and κ, to address two conceptually distinct sources of social value.

C.1.6 Decentralized equilibrium

In the decentralized economy, an entrepreneur of ability z enters if expected monopoly profit (averaged over direction draws) exceeds the entry cost. Using the certainty-equivalent approximation $\mathbb{E}[\theta_i \mid z, d^*] = (1+bz)/(2+bz)$, the entry condition $\Pi(z) \geq f w_0$ yields a cutoff $\tilde{z}^d$ and equilibrium entry m^d. This decentralized benchmark contains atomistic entrepreneurs but no portfolio intermediary. By Assumption 1, no individual entrepreneur finds it profitable to purchase industry-wide directional research for a single draw, so aggregate atomistic directional research is $\kappa^d = 0$ and $q^d = 0$.

C.1.7 GGF and PE: Stage 1 and Stage 2 first-order conditions

Under the symmetric assumption stated in the main text, the GGF and PE share the per-firm screening cost $c > 0$, the equity stake $\gamma \in (0,1)$, the directional-research technology $q(\cdot)$, and the enhancement technology $\bar{b} > b$. The single dimension on which they differ is the social-return weight λ: the GGF internalizes a fraction $\lambda \in (0,1)$ of social returns, the PE fund has $\lambda = 0$.

Stage 1: Directional research. Taking the supported set H as given, the GGF chooses κ^{GGF} to satisfy

$$\sum_{i \in H} \left[\gamma \frac{\partial \Delta \Pi_i}{\partial q} + \lambda \frac{\partial \Delta W_i}{\partial q} \right] q'(\kappa^{GGF}) = 1, \tag{A.8}$$

where $\Delta \Pi_i$ is the supported entrepreneur's profit gain and ΔW_i is the social-welfare gain from supporting i. The PE counterpart, taking H^{PE} as given, satisfies

$$\sum_{i \in H^{PE}} \gamma \frac{\partial \Delta \Pi_i}{\partial q} q'(\kappa^{PE}) = 1. \tag{A.9}$$

The two FOCs differ only in the additional $\lambda \sum \partial \Delta W_i / \partial q$ term that the GGF internalizes. The marginal cost of κ is identical at unity.

Stage 2: Screening with symmetric equity and cost. Given the period-1 directional outcome q from Stage 1, the GGF supports entrepreneur i if and only if

$$\gamma \Delta \Pi(z_i; q) + \lambda \, \Delta W(z_i; q) \; \geq \; c, \tag{A.10}$$

where $\Delta \Pi(z_i; q)$ reflects both the capability boost ($b \to \bar{b}$) and the directional advantage (probability $q + (1-q)/K$ of pursuing d^*), and $\Delta W(z_i; q)$ is the social-welfare gain. A PE fund with the *same* (c, γ) supports i if and only if

$$\gamma \Delta \Pi(z_i; q) \; \geq \; c. \tag{A.11}$$

Equivalently, in rearranged form, the GGF requires $\Delta \Pi(z) \geq [c - \lambda \Delta W(z)]/\gamma$ while the PE fund requires $\Delta \Pi(z) \geq c/\gamma$. The two conditions differ *only* through the social-welfare term $\lambda \Delta W$.

C.1.8 Increasing differences and the screening threshold

Assumption 4 (Increasing Differences in Ability)**.** *For each q, the private and social gains from support, $\Delta \Pi(z;q)$ and $\Delta W(z;q)$, are continuous and strictly increasing in z on $[\underline{z}, \bar{z}]$. The crossing conditions $B^{PE}(\underline{z};q) < 0 < B^{PE}(\bar{z};q)$ and $B^{GGF}(\underline{z};q,\lambda) < 0 < B^{GGF}(\bar{z};q,\lambda)$ hold for the relevant*

parameter range.

Assumption 4 is a single-crossing condition: ability and support are complements. It is not implied by the simple certainty-equivalent profit expression in every parameter region. For example, if the supported and unsupported success probabilities are already close to one at the top of the ability distribution, the private profit gain can become locally flat. We therefore impose increasing differences directly rather than deriving it from an approximation that does not hold globally. This assumption underpins Lemma 2 in the main text.

C.1.9 Joint equilibrium of screening and directional research

The Stage 1 problem depends on $|H|$ (a Stage 2 outcome), and the Stage 2 problem depends on $q(\kappa)$ (a Stage 1 outcome). To break this circularity rigorously, we define the GGF equilibrium as a fixed point.

Assumption 5 (Stable Fund Equilibrium)**.** *For both PE and GGF, the composite best-response map between the support threshold and directional research has a unique interior stable fixed point. Locally, the absolute slope of the composite map is strictly below one at that fixed point.*

Assumption 5 is a regularity condition on the joint fund problem. It rules out multiple self-fulfilling portfolio-size equilibria, which can arise in models where a larger portfolio raises the value of information and better information lowers the support threshold. We impose uniqueness as a stability condition rather than deriving it from concavity or monotonicity alone.

Define the GGF best-response maps

$$\Psi^{GGF} : \tilde{z} \mapsto \arg\max_{\kappa \geq 0} \sum_{i \in H(\tilde{z})} \left[\gamma \Pi_i(q(\kappa)) + \lambda W_i(q(\kappa))\right] - \kappa, \qquad \Phi^{GGF} : \kappa \mapsto \tilde{z}^{GGF}(q(\kappa), \lambda).$$

Both are continuous and weakly monotone in their arguments. Under Assumption 5, there exists a unique stable fixed point $(\kappa^{GGF}, \tilde{z}^{GGF})$ of $T^{GGF} \equiv \Phi^{GGF} \circ \Psi^{GGF}$. The PE problem yields a unique stable fixed point $(\kappa^{PE}, \tilde{z}^{PE})$ analogously (with $\lambda = 0$). The proof that these fixed points satisfy the ordering reported in Lemma 3 of the main text is in Appendix C.3.4.

C.1.10 GGF frontier and Quality Filtering channels

Under the GGF regime, let $H = \{i : z_i \geq \tilde{z}^{GGF}\}$ denote supported entrepreneurs and N the rest. Conditional on the realized directional allocation D^*, the expected frontier is

$$\mathbb{E}[\hat{\theta}^{GGF} \mid D^*] = 1 - \frac{1}{1 + m + \bar{b}\sum_{i \in H \cap D^*} z_i + \delta\bar{b}\sum_{i \in H \setminus D^*} z_i + b\sum_{j \in N \cap D^*} z_j + \delta b \sum_{j \in N \setminus D^*} z_j}. \quad \text{(A.12)}$$

Because $q^{GGF} > q^{PE}$, supported firms are more likely to be in D^*. Taking expectations over directions, the unconditional GGF frontier exceeds the PE and no-intervention frontiers.

C.1.11 Budget identities for the four policy instruments

Each instrument allocates B subject to its own budget identity:

- **Subsidy.** A per-entrant subsidy s raises entry by lowering the effective cost from fw_0 to $fw_0 - s$. The budget identity is $B = s \cdot m^{sub}(s)$.
- **Tax incentive.** A reduced tax rate raises post-discovery profits from $\Pi(z)$ to $(1+\rho)\Pi(z)$ at fiscal cost $B = \rho \cdot \mathbb{E}[\Pi] \cdot m^{tax}(\rho)$.
- **Patent extension.** Lengthening monopoly duration T raises private profits at a welfare cost from delayed period-2 output. We treat $B = 0$ in fiscal terms, but the welfare cost is captured by a deadweight term $C'_{pat}(B) > 0$.
- **GGF.** Spending is split between per-firm support and directional research: $B = c \cdot |H^{GGF}| + \kappa^{GGF}$.

The Unique-Coverage proposition in the main text (Proposition 5) compares these instruments on the basis of which margins (m or q) their budget can move.

C.2 Policy Instruments: Budget Identities and Formal Results

This appendix collects the formal budget-equivalent framework and the two main results—Proposition 5 on unique coverage of the directional externality, and Corollary 2 on conditions for budget-equivalent GGF dominance—that the main text describes in narrative form. The proofs are in Appendices C.3.8 and C.3.9.

C.2.1 Budget-equivalent setup

Each instrument allocates a common fiscal budget $B > 0$ subject to its own budget identity:

- **Subsidy.** A per-entrant subsidy s raises entry by lowering the effective cost from fw_0 to $fw_0 - s$. The budget identity is $B = s \cdot m^{sub}(s)$.
- **Tax incentive.** A reduced tax rate ρ raises post-discovery profits from $\Pi(z)$ to $(1+\rho)\Pi(z)$ at fiscal cost $B = \rho \cdot \mathbb{E}[\Pi] \cdot m^{tax}(\rho)$.
- **Patent extension.** Lengthening monopoly duration T raises private profits at a welfare cost from delayed period-2 output. We treat $B = 0$ in fiscal terms, but the welfare cost is captured by a deadweight term $C'_{pat}(B) > 0$.
- **GGF.** Spending is split between per-firm support and directional research: $B = c \cdot |H^{GGF}| + \kappa^{GGF}$.

This formulation directly addresses the standardization concern that comparing $\partial \mathbb{E}[\hat{\theta}]/\partial \bar{b}$ against $\partial \mathbb{E}[\hat{\theta}]/\partial f$ confounds units of policy intensity. Under a budget-equivalent comparison, all instruments spend the same number of dollars, and we ask which dollars buy the largest welfare improvement.

C.2.2 Unique coverage of the directional externality

Proposition 5 (Unique Coverage of the Directional Externality)**.** *For any budget $B > 0$:*

(i) Subsidies, tax incentives, and patent extensions affect only the extensive margin: $\partial m^I/\partial B > 0$ *and* $\partial q^I/\partial B = 0$ *for* $I \in \{\textit{Subsidy, Tax, Patent}\}$*. They cannot reduce the directional wedge* κ^**.*

(ii) GGF affects both margins: $\partial m^{GGF}/\partial B > 0$ *and* $\partial q^{GGF}/\partial B > 0$*.*

Subsidies, tax incentives, and patent extensions are thus "one-knob" instruments: they can buy more entry but not better directional allocation. The reason is structural. Each operates on the individual entrepreneur's profit, whereas directional research is a portfolio-level fixed cost (Assumption 1), so an instrument that does not bundle entrepreneurs into a portfolio cannot deliver $q > 0$. Only a fund that internalizes the portfolio externality can move the directional margin, and the distinction is structural rather than quantitative: the traditional instruments cannot close the directional wedge however large their budget. The proof is in Appendix C.3.8.

C.2.3 Conditions for budget-equivalent GGF dominance

Corollary 2 (Conditions for Budget-Equivalent GGF Dominance)**.** *Let* $W_I(B)$ *denote the welfare gain from instrument* $I \in \{\textit{Subsidy, Tax, Patent, GGF}\}$ *at budget* B*. Then* $W_{GGF}(B) > W_I(B)$ *for* $I \in \{\textit{Subsidy, Tax, Patent}\}$ *whenever*

$$\underbrace{Y_2'(\Theta)\cdot\Theta_q\cdot\frac{\partial q^{GGF}}{\partial B}}_{\textit{directional value of an extra GGF dollar}} \quad > \quad \underbrace{Y_2'(\Theta)\cdot\Theta_m\cdot\left[\frac{\partial m^I}{\partial B}-\frac{\partial m^{GGF}}{\partial B}\right]}_{\textit{extensive-margin handicap of GGF, if any}}.$$

The condition is satisfied whenever the directional externality is at least as binding as the knowledge externality (i.e., $\kappa^* - \kappa^d$ *is large relative to* $m^* - m^d$ *at the policy-relevant budget).*

The corollary gives a unit-free, budget-equivalent comparison. The GGF dominates not because its marginal effect on a single parameter is mechanically larger, but because (a) it moves both equilibrium objects (m,q) rather than only m, and (b) the social value of the extra movement in q exceeds any slippage on the extensive margin. The ranking is conditional, not universal. When the directional externality is small or absent, as in a hypothetical industry where d^* is already known,

the traditional instruments can outperform the GGF because they move m more cheaply per dollar. When it is large, as in emerging industries with high R&D intensity and uncertain technological paths, the GGF is the welfare-maximizing instrument. The full proof, including the integration argument that takes the marginal inequality to the welfare-level inequality, is in Appendix C.3.9.

C.3 Proofs

C.3.1 Proof of Lemma 1

Proof. For each $i \in \{1,\dots,m\}$, the conditional CDF of θ_i on $[0,1]$ is $F^*(x;z_i) = x^{1+bz_i}$ if $i \in D^*$ and $F^0(x;z_i) = x^{1+\delta bz_i}$ if $i \notin D^*$. Conditional on D^* and $\{z_i\}$, the productivity draws $\{\theta_i\}_{i=1}^m$ are independent. Hence

$$\Pr\Big(\hat{\theta} \le x \mid D^*, \{z_i\}\Big) \;=\; \prod_{i \in D^*} x^{1+bz_i} \cdot \prod_{j \notin D^*} x^{1+\delta bz_j} \;=\; x^S,$$

where

$$S \;\equiv\; \sum_{i \in D^*} (1+bz_i) + \sum_{j \notin D^*} (1+\delta bz_j) \;=\; m + b\sum_{i \in D^*} z_i + \delta b \sum_{j \notin D^*} z_j.$$

Since $\hat{\theta} \in [0,1]$ almost surely,

$$\mathbb{E}[\hat{\theta} \mid D^*, \{z_i\}] \;=\; \int_0^1 \big(1 - \Pr(\hat{\theta} \le x)\big)\,\mathrm{d}x \;=\; \int_0^1 \big(1 - x^S\big)\mathrm{d}x \;=\; 1 - \frac{1}{1+S},$$

which is (A.3).

Marginal value of directional reallocation. Reallocating one entrepreneur of ability z from a wrong direction to d^* replaces δbz in S by bz, raising S by $(1-\delta)bz$. The induced change in the expected frontier is

$$\frac{1}{1+S} - \frac{1}{1+S+(1-\delta)bz} \;=\; \frac{(1-\delta)bz}{(1+S)\big(1+S+(1-\delta)bz\big)} \;>\; 0,$$

which is strictly increasing in z on $[\underline{z}, \bar{z}]$. □

C.3.2 Proof of Proposition 1

Proof. We establish each part separately.

Part (i): Existence and uniqueness of $\hat{T}$. The decentralized cutoff $\tilde{z}^d(T)$ solves $\Pi(z;T) = fw_0$, where the ex-ante expected monopoly profit is

$$\Pi(z;T) \;=\; T\cdot\left(p-\frac{w_1}{A\,\overline{\theta}(z)}\right), \qquad \overline{\theta}(z)\equiv \tfrac{1}{K}\tfrac{1+bz}{2+bz}+\tfrac{K-1}{K}\tfrac{1+\delta bz}{2+\delta bz}.$$

The function $T\mapsto\Pi(z;T)$ is strictly increasing in T for any $z>0$, so $T\mapsto\tilde{z}^d(T)$ is continuous and strictly decreasing. Equilibrium entry $m^d(T)\equiv\mu(\{z:z\geq\tilde{z}^d(T)\})$, where μ is the population measure on ability, is therefore continuous and strictly increasing in T.

The social-planner solution (m^*,κ^*) does not depend on T because the planner values total period-2 output $Y_2(\Theta(m,q(\kappa)))$, which is independent of monopoly duration. Define the function

$$G(T) \;\equiv\; m^d(T)-m^*.$$

Because m^d is continuous and strictly increasing in T, with $\lim_{T\downarrow 0}m^d(T)=0<m^*$ and $\lim_{T\uparrow\infty}m^d(T)>m^*$ (full entry exceeds the planner's optimum once private rents are unbounded), the intermediate-value theorem yields a unique $\hat{T}\in(0,\infty)$ such that $G(\hat{T})=0$. For any $T<\hat{T}$, $G(T)<0$, i.e., $m^d(T)<m^*$.

Part (ii): Directional externality. In the decentralized benchmark, entrepreneurs are atomistic and do not operate portfolios. The planner's first-order condition (A.7) requires

$$Y_2'(\Theta^*)\cdot\Theta_q(m^*,q^*)\cdot q'(\kappa^*)=1.$$

By Assumption 2, the planner's marginal portfolio value of directional research at $\kappa=0$ exceeds one and tends to zero as $\kappa\to\infty$. Since q is strictly concave, the planner has a unique interior optimum $\kappa^*>0$.

For an individual entrepreneur controlling a single firm, the marginal benefit of choosing any

positive amount of directional research is bounded by

$$\frac{\partial \Pi_i}{\partial q} \cdot q'(\kappa) \;=\; \left[\Pi^*(z_i) - \Pi^0(z_i)\right] \cdot \frac{K-1}{K} \cdot q'(\kappa) \;\leq\; \sup_z [\Pi^*(z) - \Pi^0(z)] \cdot \frac{K-1}{K} \cdot q'(\kappa).$$

Because $q'(0)$ is finite and q' is decreasing, Assumption 1 implies this marginal benefit is strictly less than 1 for all $\kappa \geq 0$. Hence the individual entrepreneur's optimum is zero. Since every entrepreneur faces the same single-firm bound and there is no intermediary in this benchmark, aggregate atomistic directional research is $\kappa^d = 0$.

Furthermore, even if an entrepreneur could discover d^*, the information would diffuse through observed outcomes via the inference rule of Assumption 3 once a firm on d^* succeeds. The discovering entrepreneur bears the full cost but does not capture the sector-wide benefit, reinforcing $\kappa^d = 0$. □

C.3.3 Proof of Lemma 2

Proof. We proceed in four steps.

Step 1: Monotonicity of private and social gains. By Assumption 4, $\Delta\Pi(z;q)$ and $\Delta W(z;q)$ are continuous and strictly increasing in z for each q. We need this assumption because the simple certainty-equivalent profit expression does not globally imply increasing private gains from support. For some parameters, the private gain from raising b becomes locally flat for very high-ability entrepreneurs. The threshold result therefore rests on the single-crossing condition itself, not on a derivative calculation that fails globally.

Step 2: Existence and uniqueness of thresholds. Define

$$B^{GGF}(z;q,\lambda) \equiv \gamma\Delta\Pi(z;q) + \lambda\Delta W(z;q) - c, \qquad B^{PE}(z;q) \equiv \gamma\Delta\Pi(z;q) - c.$$

Both functions are continuous on $[\underline{z},\bar{z}]$ and strictly increasing in z. The crossing conditions in Assumption 4 imply that the intermediate-value theorem yields unique thresholds $\tilde{z}^{GGF}(q,\lambda)$ and $\tilde{z}^{PE}(q)$ at which $B^I = 0$.

Step 3: GGF vs. PE threshold ranking. For any $z \in [\underline{z}, \bar{z}]$,

$$B^{GGF}(z;q,\lambda) - B^{PE}(z;q) \;=\; \lambda\, \Delta W(z;q).$$

Since $\Delta W(z;q) > 0$ for all relevant z, $B^{GGF}(z;q,\lambda) > B^{PE}(z;q)$ if and only if $\lambda > 0$. The threshold is the smallest z at which $B^I = 0$. Because B^{GGF} is uniformly strictly larger than B^{PE} when $\lambda > 0$, and both are strictly increasing, $\tilde{z}^{GGF}(q,\lambda) < \tilde{z}^{PE}(q)$. When $\lambda = 0$, the two thresholds coincide. □

C.3.4 Proof of Lemma 3

Proof. **Setup.** Let $\mathscr{Z} \equiv [\underline{z}, \bar{z}]$ and assume the population distribution of ability is non-atomic on $\mathscr{Z}$, so $|H(\tilde{z})| \equiv \mu(\{z : z \geq \tilde{z}\})$ is continuous and strictly decreasing in $\tilde{z}$.

Best-response maps. Define

$$\Psi^{GGF}(\tilde{z}) \;\equiv\; \arg\max_{\kappa \geq 0} \sum_{i \in H(\tilde{z})} \left[\gamma \Pi_i(q(\kappa)) + \lambda W_i(q(\kappa))\right] - \kappa,$$

the GGF's optimal κ given a conjectured threshold $\tilde{z}$. By strict concavity of $q(\kappa)$ and (A.8), Ψ^{GGF} is single-valued and continuous in $\tilde{z}$. A higher $\tilde{z}$ shrinks H, lowering the marginal benefit of κ, so Ψ^{GGF} is weakly decreasing.

Similarly, define

$$\Phi^{GGF}(\kappa) \;\equiv\; \tilde{z}^{GGF}(q(\kappa),\lambda),$$

the GGF's optimal threshold given a conjectured Stage 1 outcome κ. By Lemma 2, Φ^{GGF} is single-valued and continuous in κ. A higher κ raises q, which raises both $\Delta\Pi$ and ΔW, lowering $\tilde{z}^{GGF}$. Hence Φ^{GGF} is weakly decreasing in κ.

Existence of fixed point. Consider the composite map $T^{GGF}(\tilde{z}) \equiv \Phi^{GGF}(\Psi^{GGF}(\tilde{z}))$. As the composition of two continuous, weakly decreasing maps from $\mathscr{Z}$ into $\mathscr{Z}$, T^{GGF} is continuous and weakly increasing on the compact set $\mathscr{Z}$. By the Brouwer fixed-point theorem, T^{GGF} has at least

one fixed point $\tilde{z}^{GGF} \in \mathscr{Z}$. Set $\kappa^{GGF} \equiv \Psi^{GGF}(\tilde{z}^{GGF})$.

Uniqueness and stability. Monotonicity and concavity alone do not imply that the composite map is a contraction. We therefore invoke Assumption 5, which imposes a unique interior stable fixed point and a local slope with absolute value below one. This rules out multiple self-fulfilling portfolio-size equilibria without incorrectly deriving contraction from concavity.

The PE problem is identical with $\lambda = 0$, so the analogous map T^{PE} has a unique fixed point $(\kappa^{PE}, \tilde{z}^{PE})$.

Comparison: $\tilde{z}^{GGF} < \tilde{z}^{PE}$ and $\kappa^{GGF} > \kappa^{PE}$. For any $\kappa > 0$, Lemma 2 gives $\Phi^{GGF}(\kappa) < \Phi^{PE}(\kappa)$. Hence $T^{GGF}(\tilde{z}) < T^{PE}(\tilde{z})$ for any $\tilde{z} \in (\underline{z}, \bar{z})$. Combined with monotonicity and the stable fixed-point selection in Assumption 5, the fixed point of T^{GGF} is strictly smaller than that of T^{PE}:

$$\tilde{z}^{GGF} < \tilde{z}^{PE} \quad \text{and hence} \quad |H^{GGF}| > |H^{PE}|.$$

Plugging back into the FOCs (A.8) and (A.9), the GGF's marginal benefit at any κ is uniformly strictly larger than the PE's: (a) $\lambda > 0$ adds the strictly positive term $\sum_{i \in H^{GGF}} \partial \Delta W_i / \partial q$; (b) $|H^{GGF}| > |H^{PE}|$ scales up the existing $\partial \Delta \Pi_i / \partial q$ terms. By strict concavity of $q(\kappa)$, the unique solution of (A.8) satisfies $\kappa^{GGF} > \kappa^{PE}$.

Comparison with planner: $\kappa^{GGF} < \kappa^{*}$. The planner's FOC (A.7) can be written as

$$\sum_{i=1}^{m^*} \left[Y_2'(\Theta^*) \cdot \frac{\partial \Theta}{\partial q}\bigg|_i \right] q'(\kappa^*) = 1.$$

Two differences make the planner's marginal benefit at any κ strictly larger than the GGF's: (a) the planner sums over all m^* entrepreneurs, while the GGF sums over the strict subset H^{GGF}; (b) the planner values the entire social return $Y_2'(\Theta)$, while the GGF captures only the fraction $\gamma + \lambda < 1$. Both differences raise the planner's marginal benefit curve uniformly above the GGF's. By strict concavity of $q(\kappa)$, $\kappa^* > \kappa^{GGF}$.

Lower bound: $\kappa^{PE} > \kappa^{d} = 0$. By Assumption 2, a PE fund's marginal portfolio value of directional research at $\kappa = 0$ exceeds one. Hence its interior optimum satisfies $\kappa^{PE} > 0$, while

Assumption 1 gives $\kappa^d = 0$ for an individual entrepreneur. □

C.3.5 Proof of Proposition 2

Proof. We prove the result for potential productivities before the endogenous production cutoff. This avoids any claim about the sign of the period-1 wage response, which is ambiguous under unit output per firm because labor demand per firm equals $1/(A\theta_i)$ and is decreasing in productivity.

Channel 1: Entry selection. The GGF supports the entrepreneurs in $[\tilde{z}^{GGF}, \tilde{z}^{PE})$ identified in Corollary 1. By Assumption 4, these entrepreneurs have positive support gains and are above the GGF threshold. Thus the supported part of the entrant pool is selected on higher ability relative to the marginal unsupported pool.

Channel 2: Capability improvement. For any supported entrepreneur with ability $z > 0$, $\bar{b} > b$ implies that on each direction the productivity CDF $x^{1+\bar{b}z}$ lies strictly below x^{1+bz} on $(0,1)$:

$$x^{1+\bar{b}z} < x^{1+bz} \quad \text{for all } x \in (0,1),$$

since $1+\bar{b}z > 1+bz$. This is FOSD of individual productivity draws among supported firms.

Channel 3: Directional guidance. Supported entrepreneurs pursue d^* with probability $q^{GGF} + (1-q^{GGF})/K > 1/K$. On d^*, the CDF is $x^{1+\bar{b}z}$; on a wrong direction, it is $x^{1+\delta\bar{b}z}$ with $\delta < 1$. The mixture CDF for a supported entrepreneur is

$$G_i(x;q) \;=\; \left[q + \tfrac{1-q}{K}\right] x^{1+\bar{b}z_i} + \tfrac{(1-q)(K-1)}{K} x^{1+\delta\bar{b}z_i}.$$

Because $x^{1+\bar{b}z_i} < x^{1+\delta\bar{b}z_i}$ for all $x \in (0,1)$ when $\delta < 1$, $G_i(x;q)$ is strictly decreasing in q on $(0,1)$. A higher q^{GGF} thus FOSD-shifts the conditional productivity distribution upward.

Combining. Let Φ_i^{GGF} and Φ_i^{NoGGF} denote the CDFs of θ_i for entrepreneur i under each regime. Channels 2 and 3 deliver $\Phi_i^{GGF} \leq \Phi_i^{NoGGF}$ for every supported i, with strict inequality at all interior x. Channel 1 increases the average ability of the supported entrant pool, which by the parameterization $F^*(x;z) = x^{1+bz}$ further FOSD-shifts the conditional productivity distribution.

Aggregating across entrants, the unconditional CDF satisfies

$$\mathscr{F}_{\text{GGF}}(\theta) < \mathscr{F}_{\text{NoGGF}}(\theta) \quad \text{for all } \theta \in (0,1),$$

which is FOSD. □

C.3.6 Proof of Proposition 3

Proof. **Setup.** Let $m = m^I$ be the number of period-1 entrants under regime $I \in \{GGF, NoGGF\}$. Under no GGF, each entrepreneur draws d_i uniformly on $\{1,\dots,K\}$, so $\{n_k\}_{k=1}^K$ follows a multinomial$(m, 1/K, \dots,$ distribution. The expected Herfindahl index is

$$\mathbb{E}[\mathscr{S}^{NoGGF}] = \mathbb{E}\left[\sum_{k=1}^{K}(n_k/m)^2\right] = \frac{1}{K} + \frac{K-1}{Km}.$$

Under GGF, each supported entrepreneur ($i \in H$, with $|H|$ deterministic at the joint equilibrium) chooses d^* with probability $p_H = q^{GGF} + (1-q^{GGF})/K$ and each unsupported entrepreneur chooses uniformly. The expected number on d^* is

$$\begin{aligned}\mathbb{E}[n_{d^*}^{GGF}] &= |H| p_H + (m - |H|) \cdot \tfrac{1}{K} \\ &= |H| q^{GGF} \cdot \tfrac{K-1}{K} + \tfrac{m}{K}.\end{aligned}$$

This exceeds the no-GGF expectation m/K by $\Delta d \equiv |H| q^{GGF} \cdot (K-1)/K > 0$.

Part (i): Short-run condition. Let $m_0 = m^{NoGGF}$, $m_1 = m^{GGF}$, and $x = \Delta d / m_1$. Under no GGF, by the standard moment formula for the multinomial Herfindahl,

$$\mathbb{E}[\mathscr{S}_1^{NoGGF}] = \frac{1}{K} + \frac{K-1}{Km_0}.$$

Under GGF, the probability of choosing d^* is $p_* = 1/K + x$. The remaining probability mass is spread symmetrically across the $K-1$ non-viable directions, so each other direction has probability

$p_o = 1/K - x/(K-1)$. The population Herfindahl term is therefore

$$A(x) \equiv p_*^2 + (K-1)p_o^2 = \left(\frac{1}{K} + x\right)^2 + (K-1)\left(\frac{1}{K} - \frac{x}{K-1}\right)^2 = \frac{1}{K} + \frac{K}{K-1}x^2.$$

For a multinomial sample of size m_1 with cell probabilities $\{p_k\}$, $\mathbb{E}[\sum_k (n_k/m_1)^2] = \sum_k p_k^2 + (1 - \sum_k p_k^2)/m_1$. Hence

$$\mathbb{E}[\mathscr{S}_1^{GGF}] = A(x) + \frac{1 - A(x)}{m_1}.$$

Therefore $\mathbb{E}[\mathscr{S}_1^{GGF}] < \mathbb{E}[\mathscr{S}_1^{NoGGF}]$ if and only if

$$A(x) + \frac{1 - A(x)}{m_1} < \frac{1}{K} + \frac{K-1}{Km_0},$$

which is the entry-expansion condition referenced in Proposition 3(i) of the main text. The coefficient on x^2 is $K/(K-1)$, not $(K-1)/K$.

Part (ii): Medium-run convergence. Define the diffusion event

$$R \;\equiv\; \{\hat{d} = d^*\} \;=\; \left\{\arg\max_d \max_{i:\, d_i = d} \theta_i = d^*\right\},$$

with $\hat{d}$ given by Assumption 3. Under either regime,

$$\Pr(R \mid I) \;=\; \Pr\Big(\max_{i:\, d_i = d^*} \theta_i \;>\; \max_{j:\, d_j \neq d^*} \theta_j \;\Big|\; I\Big).$$

By Lemma 1, the maximum order statistic on d^* has CDF

$$H_{d^*}^I(x) \;=\; x^{|D^* \cap H| \cdot (1 + \bar{b}\,\bar{z}_{d^*}^H) + |D^* \setminus H| \cdot (1 + b\,\bar{z}_{d^*}^N)},$$

where $\bar{z}$ denotes average ability within the indicated subset. Under GGF, $\bar{b} > b$ enters and $\mathbb{E}|D^* \cap H| > 0$, both of which strictly raise the exponent and hence FOSD-shift $H_{d^*}^{GGF}$ above $H_{d^*}^{NoGGF}$. Symmetrically, the maximum order statistic on directions other than d^* has the same exponent

under both regimes (since unsupported entrepreneurs are present in both regimes and supported entrepreneurs preferentially choose d^* under GGF). Hence the probability that the global maximum is on d^* satisfies

$$\Pr(R \mid GGF) \;>\; \Pr(R \mid NoGGF).$$

In period 2, conditional on R, all new entrants and pivoting firms choose d^* (Assumption 3). Patent similarity converges to $\mathscr{S}^{\text{conv}} \approx 1$ as new entrants flood d^*. Conditional on $\neg R$, agents follow an incorrect direction, and the resulting Herfindahl is bounded by some $\mathscr{S}^{\text{diff}} < \mathscr{S}^{\text{conv}}$. Hence

$$\mathbb{E}[\mathscr{S}_2^I] \;=\; \Pr(R \mid I) \cdot \mathscr{S}^{\text{conv}} \;+\; (1 - \Pr(R \mid I)) \cdot \mathscr{S}^{\text{diff}}.$$

Since $\mathscr{S}^{\text{conv}} > \mathscr{S}^{\text{diff}}$ and $\Pr(R \mid GGF) > \Pr(R \mid NoGGF)$, $\mathbb{E}[\mathscr{S}_2^{GGF}] > \mathbb{E}[\mathscr{S}_2^{NoGGF}]$.

Part (iii): Monotonicity. Differentiating in q^{GGF},

$$\frac{\partial \mathbb{E}[\mathscr{S}_2^{GGF}]}{\partial q^{GGF}} \;=\; \left(\mathscr{S}^{\text{conv}} - \mathscr{S}^{\text{diff}}\right) \cdot \frac{\partial \Pr(R \mid GGF)}{\partial q^{GGF}}.$$

A higher q^{GGF} raises the expected number of high-$\bar{b}$ entrepreneurs on d^*, which raises the exponent of $H_{d^*}^{GGF}$ uniformly, hence raises $\Pr(R \mid GGF)$. Therefore $\partial \mathbb{E}[\mathscr{S}_2^{GGF}]/\partial q^{GGF} > 0$. By the chain rule with $q'(\kappa) > 0$,

$$\frac{\partial \mathbb{E}[\mathscr{S}_2^{GGF}]}{\partial \kappa^{GGF}} = \frac{\partial \mathbb{E}[\mathscr{S}_2^{GGF}]}{\partial q^{GGF}} \cdot q'(\kappa^{GGF}) > 0.$$

□

C.3.7 Proof of Prediction 1

Proof. **Dynamic extension.** Embed the static model in a sequence of innovation cycles indexed by $t = 1, 2, \ldots$. At the end of each cycle, all agents observe the realized frontier $\hat{\theta}_t$ and the realized patent similarity $\mathscr{S}_t$. Let M_{t+1} denote private VC capital deployed in cycle $t+1$. Assume the law of motion

$$M_{t+1} = \Phi(\hat{\theta}_t, \mathscr{S}_t), \qquad \Phi_{\hat{\theta}} > 0, \quad \Phi_{\mathscr{S}} > 0,$$

where Φ is continuously differentiable. The first inequality reflects that private investors deploy more capital when demonstrated productivity is higher. The second reflects that lower technological uncertainty (higher concentration around the inferred direction) attracts more private capital. We treat Φ as a primitive of investor preferences over expected return and risk.

Short run ($t = 1$). By Proposition 2, $\hat{\theta}_1^{GGF} > \hat{\theta}_1^{NoGGF}$. By Proposition 3(i), $\mathscr{S}_1^{GGF} \leq \mathscr{S}_1^{NoGGF}$ may hold under the entry-expansion condition $\Delta d < \Delta m$. Hence

$$\Phi(\hat{\theta}_1^{GGF}, \mathscr{S}_1^{GGF}) - \Phi(\hat{\theta}_1^{NoGGF}, \mathscr{S}_1^{NoGGF}) \;=\; \underbrace{\Phi_{\hat{\theta}} \cdot \Delta\hat{\theta}_1}_{>0} + \underbrace{\Phi_{\mathscr{S}} \cdot \Delta\mathscr{S}_1}_{\leq 0},$$

which is strictly positive whenever

$$\Phi_{\hat{\theta}} \cdot \Delta\hat{\theta}_1 \;>\; \Phi_{\mathscr{S}} \cdot |\Delta\mathscr{S}_1|.$$

This is the productivity-only channel of crowd-in. The condition is satisfied when private investors weight demonstrated returns more heavily than technological consensus, an empirically plausible feature of venture-capital decisions.

Medium run ($t \geq 2$). By Proposition 3(ii), $\mathscr{S}_t^{GGF} > \mathscr{S}_t^{NoGGF}$ for $t \geq 2$. The frontier advantage $\hat{\theta}_t^{GGF} > \hat{\theta}_t^{NoGGF}$ is preserved by the persistence of $\bar{b} > b$ and the GGF's continued $\kappa^{GGF} > 0$. Both arguments of Φ are higher under GGF, so

$$M_{t+1}^{GGF} = \Phi(\hat{\theta}_t^{GGF}, \mathscr{S}_t^{GGF}) > \Phi(\hat{\theta}_t^{NoGGF}, \mathscr{S}_t^{NoGGF}) = M_{t+1}^{NoGGF} \quad \text{for all } t \geq 2,$$

unambiguously. By induction on t, the gap $M_t^{GGF} - M_t^{NoGGF}$ widens as the directional channel reinforces the productivity channel.

Distinct timing. The productivity channel contributes to crowd-in from $t = 1$ onward, while the uncertainty-reduction channel contributes only from $t = 2$ onward. The model thus generates a two-channel crowd-in pattern with distinct horizons. □

C.3.8 Proof of Proposition 5

Proof. **Subsidies.** A per-entrant subsidy $s \geq 0$ enters the entry condition as $\Pi(z; w_0) \geq f w_0 - s$. The cutoff $\tilde{z}^{sub}(s)$ is continuous and strictly decreasing in s, so $m^{sub}(s)$ is strictly increasing. The subsidy is paid uniformly to all entrants regardless of d_i, so it does not enter the directional-research FOC (A.8) (which is unavailable to atomistic entrepreneurs by Assumption 1); $q^{sub} = q^d = 0$ is unchanged. Differentiating the budget identity $B = s \cdot m^{sub}(s)$ in B yields $\partial m / \partial B > 0$ and $\partial q^{sub} / \partial B = 0$.

Tax incentives. A tax cut $\rho \geq 0$ raises post-discovery profits to $(1+\rho)\Pi(z)$. Entry rises (the cutoff falls), but the fund-versus-self trade-off underlying directional research is unchanged: under Assumption 1, no individual entrepreneur invests in κ, regardless of the level of after-tax profits. Hence $q^{tax} = 0$ and $\partial q^{tax} / \partial B = 0$.

Patent extensions. Lengthening monopoly duration T raises $\Pi(z; T)$ proportionally. Entry rises, but again no individual entrepreneur invests in κ, so $q^{pat} = 0$ and $\partial q^{pat} / \partial B = 0$. The fiscal cost is zero, but patent extensions induce a deadweight cost $C'_{pat}(B) > 0$ from delayed period-2 diffusion (longer period-1 monopoly delays the public-good revelation of $\hat{\theta}$).

GGF. The GGF's budget is split between $c \cdot |H^{GGF}|$ and κ^{GGF}. At the joint equilibrium of Lemma 3, both margins respond strictly positively to additional B:

- $\partial m^{GGF} / \partial B > 0$: increasing B permits the GGF to lower its threshold $\tilde{z}^{GGF}$ (by Lemma 2, $\tilde{z}^{GGF}$ is strictly increasing in c, hence strictly decreasing in B holding κ fixed), expanding $|H^{GGF}|$ and hence $m^{GGF} = |H^{GGF}| + |N|$.

- $\partial q^{GGF} / \partial B > 0$: increasing B allows higher κ^{GGF}. By (A.8) and strict concavity of $q(\kappa)$, a marginal increase in available κ raises the equilibrium $q^{GGF} = q(\kappa^{GGF})$.

Hence GGF is the unique instrument among the considered set with $\partial q / \partial B > 0$. Equivalently, only GGF can reduce the directional wedge $\kappa^* - \kappa^d = \kappa^*$. □

C.3.9 Proof of Corollary 2

Proof. The welfare functional under instrument I at budget B is

$$W_I(B) \;=\; Y_2\big(\Theta(m^I(B), q^I(B))\big) - C_I(B),$$

where C_I collects the deadweight cost specific to I (e.g., distortion from raising effective wages under subsidies, delayed diffusion under patents). Differentiating,

$$\frac{\partial W_I}{\partial B} = Y_2'(\Theta) \cdot \left[\Theta_m \cdot \frac{\partial m^I}{\partial B} + \Theta_q \cdot \frac{\partial q^I}{\partial B}\right] - C_I'(B).$$

For $I \in \{\text{Subsidy, Tax, Patent}\}$, Proposition 5 gives $\partial q^I/\partial B = 0$, so

$$\frac{\partial W_I}{\partial B} = Y_2'(\Theta) \cdot \Theta_m \cdot \frac{\partial m^I}{\partial B} - C_I'(B).$$

For $I = GGF$,

$$\frac{\partial W_{GGF}}{\partial B} = Y_2'(\Theta) \cdot \Theta_m \cdot \frac{\partial m^{GGF}}{\partial B} + Y_2'(\Theta) \cdot \Theta_q \cdot \frac{\partial q^{GGF}}{\partial B} - C_{GGF}'(B).$$

Subtracting,

$$\frac{\partial (W_{GGF} - W_I)}{\partial B} = Y_2' \cdot \Theta_q \cdot \frac{\partial q^{GGF}}{\partial B} \;-\; Y_2' \cdot \Theta_m \cdot \left[\frac{\partial m^I}{\partial B} - \frac{\partial m^{GGF}}{\partial B}\right] \;+\; \big(C_I' - C_{GGF}'\big).$$

The first term is the directional-margin gain (uniquely accruing to GGF). The second is the extensive-margin handicap (potentially negative if traditional instruments are more efficient on the entry margin). Dropping the deadweight-cost difference (which is non-negative under standard regularity assumptions for the considered instruments), the explicit condition referenced in the main-text

statement of Corollary 2 is

$$\underbrace{Y_2'(\Theta)\cdot\Theta_q\cdot\frac{\partial q^{GGF}}{\partial B}}_{\text{directional value of an extra GGF dollar}} > \underbrace{Y_2'(\Theta)\cdot\Theta_m\cdot\left[\frac{\partial m^I}{\partial B}-\frac{\partial m^{GGF}}{\partial B}\right]}_{\text{extensive-margin handicap of GGF, if any}},$$

which is sufficient for $\partial(W_{GGF}-W_I)/\partial B>0$. Integrating from $B=0$ (where $W_I(0)=W_{GGF}(0)=0$) yields $W_{GGF}(B)>W_I(B)$.

The condition is satisfied whenever the directional externality is at least as binding as the knowledge externality. Specifically, when $\kappa^*-\kappa^d=\kappa^*$ is large relative to m^*-m^d at the policy-relevant budget, Θ_q is large at the equilibrium, and the dominance condition holds for any reasonable extensive-margin handicap. □

C.3.10 Proof of Proposition 4

Proof. We use the threshold equation and the directional-research first-order condition directly. No envelope theorem is applied, because $\mathbb{E}[\hat{\theta}^{GGF}]$ is an equilibrium outcome, not the value function maximized by the GGF.

(i) Enhancement capability $\bar{b}$. The denominator of (A.12) contains the terms $\bar{b}\sum_{i\in H\cap D^*}z_i+\delta\bar{b}\sum_{i\in H\setminus D^*}z_i$. Direct differentiation in $\bar{b}$, holding the optimal H and q fixed, yields

$$\left.\frac{\partial\mathbb{E}[\hat{\theta}^{GGF}]}{\partial\bar{b}}\right|_{\text{eq}}=\frac{\sum_{i\in H\cap D^*}z_i+\delta\sum_{i\in H\setminus D^*}z_i}{(1+S^{GGF})^2}>0.$$

Moreover, under Assumption 4, a higher $\bar{b}$ weakly raises support gains and therefore weakly lowers $\tilde{z}^{GGF}$ by the implicit-function theorem applied to $B^{GGF}(\tilde{z};q,\lambda)=0$. The expansion of H and the induced increase in the marginal value of κ reinforce the direct positive sign.

(ii) Screening cost c. Differentiating the threshold condition $B^{GGF}(\tilde{z}^{GGF};q,\lambda)-c=0$ implicitly:

$$\frac{\partial\tilde{z}^{GGF}}{\partial c}=\left.\frac{1}{\partial B^{GGF}/\partial z}\right|_{z=\tilde{z}^{GGF}}>0,$$

since B^{GGF} is strictly increasing in z (Steps 1–2 of Lemma 2). A higher $\tilde{z}^{GGF}$ shrinks H, removing

high-$\bar{b}$ contributions from the denominator of (A.12). Hence $\mathrm{d}\mathbb{E}[\hat{\theta}^{GGF}]/\mathrm{d}c < 0$.

(iii) Directional research technology q. Holding H fixed, differentiating the frontier expression yields

$$\left.\frac{\partial \mathbb{E}[\hat{\theta}^{GGF}]}{\partial q}\right|_{\text{eq}} = \frac{(1-\delta)\bar{b}\sum_{i\in H} z_i \cdot \frac{K-1}{K}}{(1+S^{GGF})^2} > 0.$$

For patent similarity, see Proposition 3(iii).

(iv) Talent pool $\bar{z}$. An increase in $\bar{z}$ adds entrepreneurs with abilities in $(\bar{z}_{\text{old}}, \bar{z}_{\text{new}})$ to the population. If these abilities exceed $\tilde{z}^{GGF}$, they join H and contribute $\bar{b}z$ or $\delta\bar{b}z$ to the denominator of (A.12); if they fall below $\tilde{z}^{GGF}$, they contribute bz or δbz. In either case, the denominator strictly increases.

(v) Pool size m. Adding an entrepreneur of ability z adds 1 to the count m and contributes either $\bar{b}z$ (or its δ-discounted version) if supported, or bz otherwise, to the denominator of (A.12). The denominator strictly increases. Additionally, a larger m weakly raises $|H|$ at the stable joint equilibrium, which raises the marginal benefit of κ in (A.8) and hence weakly raises κ^{GGF} and q^{GGF}. Both direct and indirect effects raise $\mathbb{E}[\hat{\theta}^{GGF}]$. □

Figure IA1.

This graph indicates the position of provinces in Mainland China in each tercile based on the annual GDP, the number of universities, and the local average wage in 2012.

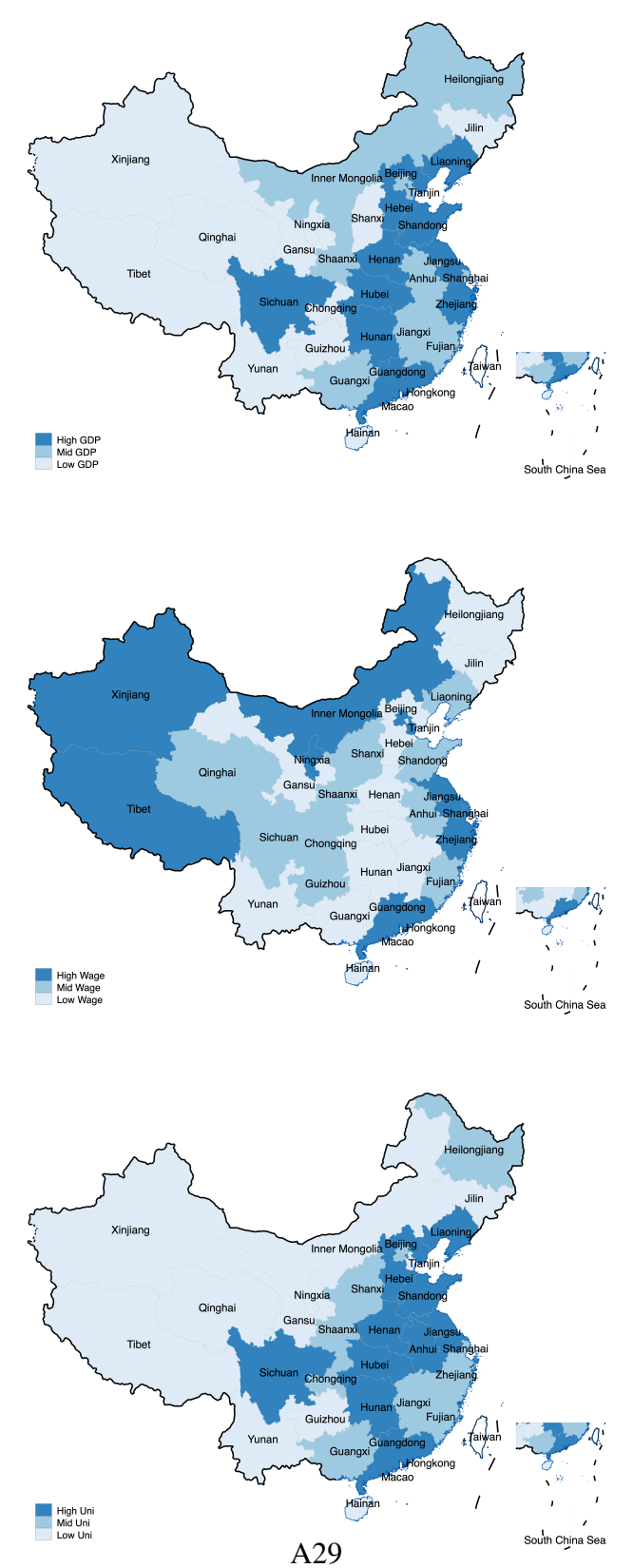


Figure IA2.

This graph indicates the position of provinces in Mainland China in each tercile based on government corruption and efficiency index taken from Nie et al. (2019)

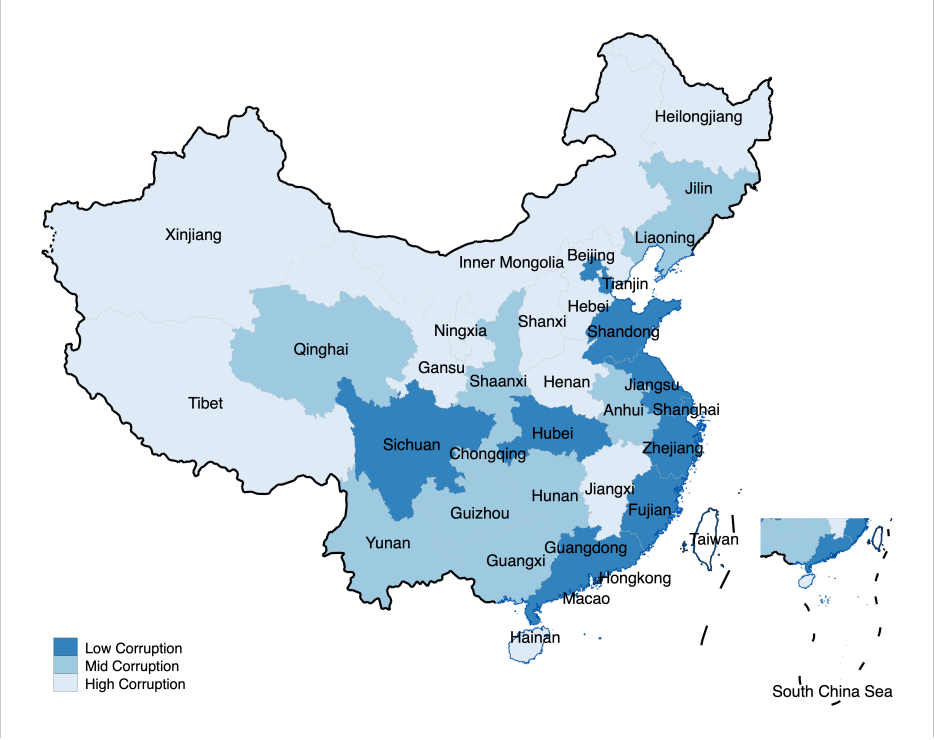




Table IA1: Robustness Checks

This table examines whether our baseline results in Table 3 are robust to alternative measures of GGF investments (Panel A) and different fixed effects (Panel B). The control variables include province-level log(GDP), average wage, and number of universities and industry-level average R&D intensity, capital intensity, ROA, log(subsidy), average age, and market-to-book value. Robust standard errors are clustered at industry × province level and are reported in parentheses. *, **, or *** indicate that the coefficient is statistically significant at the 10%, 5%, or 1% level, respectively. The details of variable definitions are in Table A1 in Appendix.

Panel A. Alternative measures of central GGF investments

	IHS(Private)		Ln(Entry)		Top_R&D		Tech_Share	
	$t+1$	$t+3$	$t+3$	$t+5$	$t+5$	$t+7$	$t+5$	$t+7$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IHS(Central_Solely)	0.837***	0.637***	0.319***	0.323***	0.071***	0.043	0.030***	0.041***
	(0.038)	(0.045)	(0.041)	(0.055)	(0.012)	(0.029)	(0.008)	(0.015)
Ln(Subsidy)	0.008	0.014	0.030***	0.036***	0.002	0.001	-0.000	0.002
	(0.007)	(0.009)	(0.012)	(0.013)	(0.003)	(0.003)	(0.002)	(0.002)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Fund Est. Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y
N	4654	4297	4599	3699	2357	1397	2357	1397
Adj. R^2	0.701	0.660	0.848	0.848	0.717	0.699	0.839	0.869

Panel B. Alternative fixed effects

	IHS(Private$_{t+1}$)		Ln(Entry$_{t+3}$)		Top_R&D$_{t+5}$		Tech_Share$_{t+5}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IHS(Central)	0.632***	0.655***	0.262***	0.183***	0.005	0.051***	0.019***	0.018***
	(0.033)	(0.030)	(0.032)	(0.032)	(0.004)	(0.008)	(0.005)	(0.004)
Ln(Subsidy)	0.008	0.003	0.018	0.025**	0.000	0.001	0.002	0.000
	(0.007)	(0.006)	(0.013)	(0.010)	(0.002)	(0.003)	(0.002)	(0.002)
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	N	Y	N	Y	N	Y	N	Y
Province FE	Y	N	Y	N	Y	N	Y	N
Year×Province FE	N	Y	N	Y	N	Y	N	Y
Year×Industry FE	Y	N	Y	N	Y	N	Y	N
N	4654	4654	4599	4599	2357	2357	2357	2357
Adj. R^2	0.740	0.734	0.862	0.913	0.876	0.709	0.839	0.910

Table IA2: SEI: Characteristics and US Sanctions

Panel A of this table compares the characteristics of SEI with those of non-SEI in the pre-SEI announcement period (i.e., before 2012). Panel B presents the probability of on the US sanction list for firms established no later than 2016 in SEI and non-SEI separately. Panel C compares the probability of facing US sanction for firms (SEI customer) whose upstream industries are in SEI with firms whose upstream industries are not. Panel D compares the probability of facing US sanction for firms (SEI supplier) whose downstream industries are in SEI with firms whose downstream industries are not. The Sample I is the Chinese tax survey sample consists of about 2.3 million firms. The Sample II contains 42,615 unique firms invested by Chinese VC/PE. *, **, or *** indicate that the difference is statistically significant at the 10%, 5%, or 1% level, respectively. The details of variable definitions are in Table A1 in Appendix.

Panel A. Pre-announcement characteristics: SEI vs. non-SEI

	N(Non-SEI)	N(SEI)	μ(Non-SEI)	μ(SEI)	Difference
R&D Intensity	299	146	0.001	0.003	-0.002***
Capital Intensity	300	146	1.496	1.073	0.423***
Market-to-Book	300	146	2.704	2.729	-0.026
ROA	300	146	0.057	0.063	-0.006
log(Subsidy)	296	146	6.165	6.218	-0.053
Age	300	146	7.926	7.529	0.398**
Observations	446				

Panel B. Probabilities of facing US sanctions: SEI vs. Non-SEI

	N(Non-SEI)	N(SEI)	μ(Non-SEI)	μ(SEI)	Difference
Sample I: Sanction Prob (%)	1,954,151	381,694	0.007	0.047	-0.039***
Sample II: Sanction Prob (%)	9,546	32,619	0.094	0.191	-0.096**

Panel C. Probabilities of facing US sanctions: SEI Customer vs. Non-SEI Customer

	N(Non-SEI Cust.)	N(SEI Cust.)	μ(Non-SEI Cust.)	μ(SEI Cust.)	Diff.
Sample I: Sanction Prob (%)	1,305,130	1,030,715	0.007	0.022	-0.015***
Sample II: Sanction Prob (%)	13,247	19,372	0.098	0.206	-0.108**

Panel D. Probabilities of facing US sanctions: SEI Supplier vs. Non-SEI Supplier

	N(Non-SEI Supp.)	N(SEI Supp.)	μ(Non-SEI Supp.)	μ(SEI Supp.)	Diff.
Sample I: Sanction Prob (%)	1,248,243	1,087,607	0.007	0.022	-0.015***
Sample II: Sanction Prob (%)	12,874	19,745	0.101	0.203	-0.102**

Table IA3: SEI vs. Non-SEI: Lead Industry

This table presents the share of lead industries in SEI and non-SEI groups from 2013 to 2016. The lead industry is defined as industries with patent self-citation ratio (i.e., number of backward citations to Chinese patents divided by the total number of backward citations for all patents in that industry) ranked in the top tercile.

	2013	2014	2015	2016
Share of Lead Industries in SEI	0	0	0.25	0.25
Share of Lead Industries in Non-SEI	0.57	0.50	0.44	0.36