Innovation Networks and R&D Allocation*

Ernest Liu†
Princeton and NBER

Song Ma‡
Yale and NBER

January 15, 2024

Abstract

We study the cross-sector allocation of R&D resources in a multisector growth model with an innovation network, where one sector's past innovations may benefit other sectors' future innovations. Theoretically, we solve for the optimal allocation of R&D resources. We show that a planner valuing long-term growth should allocate more R&D toward central sectors in the innovation network, but the incentive is muted in open economies that benefit more from foreign knowledge spillovers. We derive sufficient statistics for evaluating the welfare gains from improving R&D allocation. Empirically, we build the global innovation network based on patent citations and establish its empirical relevance for knowledge spillovers. We evaluate R&D allocative efficiency across countries using model-based sufficient statistics. Japan has the highest allocative efficiency among the advanced economies. For the U.S., improving R&D allocative efficiency could generate more than 8% welfare gains.

Keywords: Innovation Network, Resource Allocation, Growth
JEL Classification: O33, O38, F43

*Previously circulated as "Innovation Networks and Innovation Policy." We thank Chad Jones (the Editor), three anonymous referees, Philippe Aghion, Manuel Amador, Pol Antras, David Atkin, Salome Baslandze, Ariel Burstein, Lorenzo Caliendo, Vasco Carvalho, Ben Golub, Jill Grennan, Matt Jackson, Ben Jones, Chad Jones, Hugo Hopenhayn, Bill Kerr, Pete Klenow, Sam Kortum, Atif Mian, Ezra Oberfield, Bruno Pelligrino, Alessandra Peter, Stephen Redding, Peter Schott, Kjetil Storesletten, Alireza Tahbaz-Salehi, Aleh Tsyvinski, John Van Reenen, Jaume Ventura, Heidi Williams, and Kei-Mu Yi for useful feedback. We also thank many seminar and conference participants for their insights and comments. Xugan Chen, Tianyu Fan, and Daojing Zhai provided excellent research assistance.

†Department of Economics, Princeton University, and NBER, ernestliu@princeton.edu.

‡Yale School of Management, Yale University, and NBER, song.ma@yale.edu.

1 Introduction

How to foster innovation has long been a central question for economists and policy makers. The discussion has concentrated on the amount of resources invested in research and development (R&D) and the cost of under- or over-investment. But how should these R&D resources be allocated across economic sectors or technological fields? This question is important, policy relevant, yet understudied, and it is the focus of this paper. We ask: How should innovation resources be optimally allocated across sectors to take advantage of cross-sector knowledge spillovers and achieve long-term growth? For example, how many resources should an economy devote to R&D in semiconductors relative to consumer electronics, or chemistry relative to pharmaceutics? How should the optimal R&D allocations differ across countries? How are R&D resources allocated across sectors in the real world, and how much gain does it create to improve cross-sector R&D allocative efficiency?

We answer these questions both theoretically and empirically. The key novelty of our theoretical approach is that we introduce a network perspective into modeling the dynamic spillover structure of innovation. This network captures the notion that one sector's innovation activities require researchers and scientists to build on prior discoveries and knowledge, often from outside their own fields or sectors—a key feature in the innovation process. We solve for the optimal cross-sector allocation of R&D resources and derive model-implied sufficient statistics that can assess the allocative efficiency of R&D in real-world economies. The model is applied to data on more than 30 million global patents from all major economies to assess innovation resource (mis)allocation and potential welfare gains from improving allocative efficiency.

This research has two key motivations. First, cross-sector R&D allocation is an important aspect of many R&D policies, ranging from industrial policies that aim to identify and stimulate a certain set of innovative sectors, to science policy seeking to advance science and harvest long-term value. Second, the cross-sector innovation spillover structure presents a unique opportunity to combine classic endogenous growth theory, recent advances in network methods, and detailed global patent data to answer the research question.

We embed an innovation network into an otherwise canonical multisector endogenous growth model. A finite amount of R&D resources (i.e., scientists) may be deployed across sectors to innovate and improve product quality. One sector's past innovations may subsequently, over a long time path, benefit other sectors' future innovation activities by helping scientists in those sectors innovate more productively. We define the innovation network as the weighted directed graph capturing how one sector's innovation activity benefits from another's past innovation. The state variables of the economy are sectoral knowledge stocks, which reflect the accumulation of past innovations in each sector. Through dynamic spillovers across the network, the state variables form a dynamical system, in which the evolution of the knowledge stock in each sector depends endogenously on the entire history of resource allocation across all sectors of the economy. The key decision of interest is how to efficiently allocate R&D resources across sectors in the network.

We begin by modeling a closed economy. Our baseline setup adopts an analytically convenient formulation—which we relax later—where cross-sector knowledge spillovers form a log-linear dynamical system. Despite the complexity of dynamic network spillovers, we are able to explicitly solve for the optimal path of cross-sector R&D resource allocation under this formulation and express the closed-form solution in terms of consumer preferences across sectoral products and sectoral importance in stimulating future innovation through the innovation network. This solution is intuitive; it accounts for: (i) the direct effect of R&D on sectoral output, and (ii) indirect network effects on other sectors through R&D spillovers, discounting benefits that occur far in the future. The optimal R&D allocation is also related to the society's discount rate. A society valuing long-term growth (i.e., with a low discount rate) should allocate more resources toward sectors with fundamental technologies that are central in the innovation network, such as semiconductors. These are technologies that can generate widespread and long-lasting knowledge spillovers to many other sectors, directly or indirectly. By contrast, a short-termist society should allocate more R&D resources toward sectors that immediately benefit consumers but may be peripheral in the innovation network.

Formally, the contribution of each sector's R&D to economic growth is captured by the innovation network's eigenvector centrality, which we call "innovation centrality". We show the innovation centrality vector coincides with the growth-maximizing R&D allocation along a balanced growth path. The optimal R&D allocation chosen by a benevolent planner can be written as a weighted average between the innovation centrality vector and the vector representing consumer preferences over different goods. The former represents the planner's incentives to take advantage of knowledge spillovers for future growth, and the latter represents the planner's incentives to expand knowledge in ways that directly benefit the consumer. A patient planner valuing long-term growth would place a higher weight on the former.

The model also allows us to quantitatively evaluate the cross-sector allocation of R&D resources in the data and calculate the potential welfare gains from adopting the optimal R&D allocations, accounting for transitional dynamics. In consumption-equivalent terms, the welfare gains are proportional (in logs) to the inner product between the optimal R&D allocation vector and the log difference between the optimal and the actual R&D allocation vectors. Hence, this inner product—also known as the relative entropy of the actual R&D allocation from the optimal allocation—is a sufficient statistic to evaluate the potential welfare gains from improving of R&D allocation. This sufficient statistic can be calculated using data on sectoral production, the innovation network, and real-world R&D resource allocation, allowing us to quantitatively assess the

R&D allocative efficiency in the data.

Our baseline results are derived under the tractable formulation where the innovation network is exogenous and invariant to R&D allocations or the levels of knowledge stock, so that the cross-sector knowledge spillovers form a log-linear dynamical system. Under this formulation, the optimal R&D allocation is time-invariant and holds along the entire transition path; the sufficient statistic for the welfare gains from R&D reallocation also accounts for the gains along the transition path. This tractable benchmark may appear restrictive at first, but we show that our analysis holds beyond the log-linear formulation and is more general. In a richer environment with an endogenous innovation network—where cross-sector spillovers depend on the levels of knowledge stock and thus past R&D allocations—our welfare sufficient statistic is still valid to first-order around a balanced growth path; formally, it is the directional (Gateaux) derivative of welfare with respect to the allocation of R&D resources. Hence, even though the optimal R&D allocation is no longer time-invariant and instead depends on the levels of knowledge stock across sectors, our sufficient statistic serves as a first-order local approximation around a balanced growth path for the welfare gains from reallocating R&D resources; the welfare impact arising from the endogenous changes in the network (due to departure from log-linearity) is second-order in nature.

We extend our model to incorporate an important source of cross-country heterogeneity: knowledge spillovers from abroad. As we show, some countries, like the U.S. and Japan, rely more on domestic knowledge spillovers and less on foreign knowledge spillovers, while other economies benefit more from foreign spillovers particularly from the technologically advanced ones. Intuitively, from the perspective of maximizing domestic welfare, an economy receiving relatively more spillovers from abroad—and less domestically—should allocate less domestic R&D into the network-central sectors. Similarly, when domestic R&D matters less for long-run spillovers, a less efficient domestic R&D allocation is also less consequential for welfare. We demonstrate these intuitions formally by deriving, in an economy receiving foreign spillovers, both (1) the unilaterally optimal R&D allocations and (2) the sufficient statistics for the welfare gains of R&D reallocation.

The tractability of our model lends itself to a large number of theoretical extensions. We demonstrate that our baseline model, which features a simple production structure, can tractably incorporate a production network of input-output linkages. We also embed our innovation network formulation into a semi-endogenous growth setting (Jones 1995, Bloom et al. 2020, Jones 2022). We host several other extensions in the Online Appendix.

Our empirical analysis starts by constructing a global innovation network from over 36 million patents and their citations, collected from over 40 major patent authorities around the world. The data, obtained from Google Patents and originally based on the EPO worldwide bibliographic (DOCDB) data, contain patent-level information on innovations that took place in most economies between 1976 and 2020. We construct the innovation network as a weighted directed graph using sectors (and country-sectors in our open economy analysis) as nodes and citation shares from one node to another as the edge. We find innovation centrality to be highly heterogeneous across 131 3-digit international technological classes (IPCs). A handful of IPCs—such as medical science, computing, and semiconductors—are among the most central in the innovation network. Countries vary widely regarding reliance on foreign spillovers: 70% of citations made by U.S. patents are toward other U.S. patents, but most other economies—including China, South Korea, and Germany—are foreign-reliant, with domestic citation shares well below 50%. Within each country, the innovation network only weakly correlates with the input-output production network, such that there is substantial independent variation in both network structures.

To provide evidence that a sector's innovation activities benefit from past innovation in upstream sectors linked through the patent citation network, we extend Acemoglu, Akcigit, and Kerr (2016)—which analyze the U.S. domestic innovation network—using instrumental variables (IVs) and to the global setting. The IVs for past innovation are constructed based on time-varying sectoral exposure to tax-induced user cost of R&D (Wilson, 2009 and Thomson, 2017); they isolate comovements in patent output driven by knowledge spillovers and not by common shocks to connected sectors (Manski, 1993 and Bloom, Schankerman, and Van Reenen, 2013). We find evidence for directional knowledge spillovers: each sector's innovation output responds only to past upstream innovations and does not respond to past innovation from downstream sectors even though they are also connected. We also show that relative to input-output linkages, the innovation network is a significantly stronger channel through which knowledge spillovers occur.

Our main empirical application connects the model-implied optimal R&D allocation with cross-sector R&D allocation in the real world. For each country and time period, we calculate the unilaterally optimal cross-sector R&D allocation. In the U.S., sectors highly ranked in the optimal allocation are primarily those central in the innovation network, such as medical science, semiconductors, and computing devices. We find that the unilaterally optimal allocation differs significantly across countries. For instance, relative to the U.S., Germany and Japan should optimally allocate more R&D resources to vehicle-related innovation, whereas South Korea should invest more R&D in electric communication technique.

The model-implied optimal R&D allocation strongly predicts actual resource allocation in real-world economies, suggesting that our model provides a reasonable way to understand cross-sector R&D allocation. Specifically, we compare the unilaterally optimal R&D allocation against the actual R&D allocation captured using both sectoral R&D expenditure shares and patent output shares in the data. We find that, across many countries—especially for the most innovative ones—sectors that should have more R&D resources do receive more resources.

Nevertheless, the residual misalignment between the optimal and actual allocations translates into large potential welfare gains from the reallocation of R&D resources to each country's own efficient benchmark. We find Japan has the most efficient R&D allocation. From the R&D allocation in 2010, adopting the optimal allocation could lead to welfare improvements equivalent to raising consumption by 5.64% at every point in time for Japan and 8.04% in the U.S., the two economies that rely the most on domestic knowledge spillovers. Improving R&D allocation could lead to welfare gain of 5.60% in China, 4.24% in South Korea, and 4.09% in Germany. These economies' R&D allocations are less efficient than Japan's, but their domestic R&D is less welfare consequential because they benefit more from foreign spillovers. These cross-country differences in R&D allocative efficiency are qualitatively stable since the 2000s.

It is important to note that a more allocatively efficient economy is not necessarily more innovative in absolute terms. Instead, our notion of cross-sector allocative efficiency reflects the distance from an economy's actual R&D allocation in the data to this economy's own first-best, efficient benchmark. Also note that, by comparing the R&D allocations in the data to the first-best, our notion of allocative efficiency does not require that we take a stance on firms' equilibrium conduct; instead, we can directly calculate the welfare impact of reallocating R&D based on the economic environment. Finally, it is worth emphasizing again that our notion of allocative efficiency concerns the relative allocation of R&D resources across sectors and not the aggregate level of R&D.

In the final part of the paper, we discuss over- and under-allocated technology classes in the U.S. Even though providing a full policy recommendation is beyond the paper's scope, the empirical patterns are nevertheless illuminating and show our model's potential to analyze and address more detailed R&D policy issues. For example, our calculation shows that the technology class most relevant to semiconductor technologies (H01) is under-allocated in the U.S. , providing support to the recent U.S. initiatives to accelerate and catalyze the domestic semiconductor sector. We also find under-allocation of R&D in technology classes related to "green innovation" such as waste and pollution management and alternative energy.

This study relates to several strands of existing work. First, our study contributes to a long line of research on knowledge spillovers and innovation policy (Aghion et al., 2005, Bloom et al., 2013, Lucking et al., 2018, Bloom et al., 2019, Jones and Summers, forthcoming, Hopenhayn and Squintani, 2021), particularly in the context of endogenous economic growth (Jones and Williams, 1998, Ngai and Samaniego, 2011, Acemoglu et al., 2018, Akcigit and Kerr, 2018, Atkeson and Burstein, 2019, Garcia-Macia et al., 2019, Bloom et al., 2020, Akcigit et al., 2021, Cai and Tian, 2021, Koenig et al., forthcoming, Akcigit et al., 2022). We contribute to this literature by tackling a key open question: how to optimally allocate R&D resources across sectors in the presence of an innovation network with cross-sector knowledge spillovers.

Relatedly, our study connects to the literature considering cross-sector knowledge linkages, including Acemoglu et al. (2016), Cai and Li (2019), and Huang and Zenou (2020), and, in an open economy setting, Cai et al. (2022) and Guillard et al. (2021); and cross-country knowledge diffusion (Caballero and Jaffe, 1993, Jaffe et al., 1993, Eaton and Kortum, 1999, 2006, Coe and Helpman, 1995, Coe et al., 2009, Santacreu, 2015, Buera and Oberfield, 2020); see Keller (2004) and Melitz and Redding (2021) for surveys. We contribute to this literature in three ways: first, we build a new endogenous growth model explicitly considering the dynamic and cross-sector spillovers of knowledge; second, our tractable formulation enables us to derive the social optimal R&D allocation and provide simple sufficient statistics for the welfare gains of reallocating R&D optimally, and we introduce foreign-dependence as an important heterogeneity in our open economy setting; third, we construct a global innovation network using patent data around the world, which allows us to empirically study the R&D allocative efficiency in real-world economies.

We also contribute to the fast-growing literature that models network interactions in a general equilibrium setting (Carvalho, 2010, Gabaix, 2011, Acemoglu et al., 2012, Jones, 2011, 2013, Grassi, 2017, Acemoglu et al., 2015, Baqaee, 2018, Lim, 2018, Oberfield, 2018, Liu, 2019, Baqaee and Farhi, 2019, 2020, Chaney, 2018, Taschereau-Dumouchel, 2020, Kleinman et al., 2022, vom Lehn and Winberry, 2022). Particularly related are recent papers that introduce methodologies for dynamic network analysis (Liu and Tsyvinski, 2022, Kleinman et al., forthcoming) and studies on policy interventions targeting specific sectors in static production and strategic networks (Liu, 2019, Galeotti et al., 2020). Relative to this literature, our contribution is to embed an innovation network into a dynamic growth model and study the optimal allocation of R&D resources.

Finally, we contribute to the large literature of resource (mis)allocation (Restuccia and Rogerson, 2008, Hsieh and Klenow, 2009, Jones, 2013, David and Venkateswaran, 2019, Hsieh et al., 2019, Liu, 2019, Baqaee and Farhi, 2020). While this literature focuses on the static misallocation of production resources—potentially due to market distortions, such as taxes, markups and financial frictions—and primarily within-sector across firms, we study the cross-sector allocation of innovation resources, so our analysis is inherently dynamic in nature. Also note that our allocative efficiency measure does not take a stand on firms' equilibrium conduct and, instead, directly calculates the welfare impact of reallocating R&D based on the economic environment.

The rest of the paper is structured as follows. Section 2 has the model and the theoretical results. Section 3 introduces our data. Section 4 describes the global innovation network and provides evidence of its relevance for knowledge spillovers. Section 5 hosts our main empirical application, where we use the model to evaluate cross-sector R&D allocations across countries and time. Section 6 concludes. A separate Online Appendix contains the derivations of the results in the paper, theoretical extensions, and supplementary materials on data and empirical results.

2 Theory: Endogenous Growth with An Innovation Network

We study the implication of R&D resource allocation in a multisector endogenous growth model with an innovation network. We set up the baseline model in Section 2.1 and analyze the optimal allocation of R&D resources across sectors in Section 2.2. Section 2.3 shows how R&D allocation affects long-run growth along the balance growth path. Section 2.4 shows how reallocating R&D resources affects welfare, taking into account the transitional dynamics.

Our baseline model adopts a tractable log-linear formulation featuring an exogenous innovation network. Section 2.5 shows that our results are more general: in a richer environment with flexible functional forms and an endogenous innovation network, our welfare sufficient statistic is still valid to first-order around a balanced growth path.

Section 2.6 extends the baseline model with knowledge spillovers from abroad. In this setting, we derive the unilaterally optimal R&D allocation and the welfare impact of R&D reallocation.

The tractability of our model lends itself to a large number of theoretical extensions. We discuss some of these extensions in Section 2.7, such as incorporating a production network into the model (Section 2.7.1) and how to embed our innovation network formulation into a semi-endogenous growth setting (Section 2.7.2). Section 2.7.3 discuss potential inefficiencies in a stylized decentralized setting. Section 2.7.4 briefly describes other extensions in the Online Appendix.

2.1 Economic Environment of the Baseline Model

Preferences and Production Technology There is a representative consumer with log flow utility and exponential discounting at rate ρ:

$$V_t = \int_t^\infty e^{-\rho(s-t)} \ln y_s \, \mathrm{d}s. \tag{1}$$

The consumption good at each time t is a Cobb-Douglas aggregator over sectoral goods $\{y_{it}\}_{i=1}^K$:

$$y_t = \prod_{i=1}^K y_{it}^{\beta_i}, \quad \sum_{i=1}^K \beta_i = 1. \tag{2}$$

We refer to β_i as the consumption elasticity of sector i.

Each sectoral good i is produced linearly from production workers ℓ_{it}:

$$y_{it} = q_{it}^{\psi} \ell_{it}. \tag{3}$$

The sectoral productivity q_{it}^{ψ} depends on a sector's knowledge stock q_{it} at time t. The collection of sectoral knowledge stocks $\{q_{it}\}_{i=1}^K$ are the state variables of the economy. The exponent ψ parametrizes how sectoral knowledge translates into sectoral productivity.

The Innovation Process R&D expands the knowledge stock. At time t, mass s_i of R&D resources (e.g., scientists) employed in sector i generate a flow of new innovation n_{it}:

$$n_{it} = s_{it}\eta_i\chi_{it}, \qquad \chi_{it} \equiv \prod_{j=1}^{K} q_{jt}^{\omega_{ij}}, \qquad \sum_{j=1}^{K}\omega_{ij} \equiv 1. \tag{4}$$

η_i is the exogenous component of innovation productivity, and χ_{it} is the endogenous component. Importantly, χ_{it} is an aggregator over knowledge stock across all sectors. This implies that a larger knowledge stock q_j in sector j facilitates innovation production in sector i with elasticity ω_{ij}, thereby making scientists in sector i conduct R&D more productively. Our formulation thus captures cross-sector knowledge spillovers; that is, scientists stand on the shoulders of giants across all sectors of the economy. In the baseline model we assume that χ_{it} has constant returns to scale ($\sum_{j=1}^{K}\omega_{ij} = 1$) in each sector, which implies sustained and nonexplosive growth. Absent cross-sector knowledge spillovers, $\omega_{ij} = 1$ if $i = j$ and is zero otherwise.

New innovation n_{it} expands the knowledge stock according to the following law of motion:

$$\dot{q}_{it}/q_{it} = \lambda \ln\left(n_{it}/q_{it}\right). \tag{5}$$

The key distinction between n_{it} and q_{it} is that the former is a *flow* variable reflecting innovation output at time t, whereas the latter is a *stock* variable reflecting the accumulation of past innovations. The rate at which knowledge expands in sector i is increasing in the flow of innovation and decreasing in the existing knowledge stock q_{it}, capturing the notion that innovation gets harder as the knowledge stock in sector i expands. λ parametrizes the sensitivity of knowledge growth to the flow of new innovation relative to the existing stock.

Throughout the rest of the paper, we use boldface variables to denote column vectors (lowercase) and matrices (uppercase). Let $\boldsymbol{q}_t$ denote the column vector whose i-th entry is q_{it}; $\boldsymbol{q}_t$ captures the economy's state variables.

Definition 1. (**Innovation Network**) The innovation network $\boldsymbol{\Omega} \equiv [\omega_{ij}]$ is the $K \times K$ matrix whose ij-th entry is ω_{ij}.

A key object of this study, the $\boldsymbol{\Omega}$ matrix represents a weighted directed graph in which economic sectors are the graph nodes. Elements of the $\boldsymbol{\Omega}$ matrix ω_{ij} capture the elasticity to which sector i's innovation production benefits from sector j's existing knowledge stock. We refer to sector j as *upstream* to sector i and, conversely, i as *downstream* to j; this terminology captures the notion that knowledge flows from upstream sector j to downstream sector i. Absent cross-sector knowledge spillovers, $\boldsymbol{\Omega} = \boldsymbol{I}$ is the identity matrix. The construction is not limited by any specific sector definition; for instance, innovation networks can be constructed across industrial sectors, technology classes, and scientific fields. To make the network analysis interesting, we assume all sectors are strongly connected: every sector is eventually reachable from every other sector via knowledge spillovers (i.e., $\forall i, j$, $\exists k$ such that $\left[\boldsymbol{\Omega}^k\right]_{i,j} > 0$).

Resources We close the model with resource constraints. The economy is endowed with two exogenous stocks of resources: production workers of mass $\bar{\ell}$ and research scientists of mass $\bar{s}$. Workers are employed to produce sectoral goods as in (3). Scientists are employed to conduct R&D. The market clearing conditions for production workers and scientists are:

$$\sum_{i=1}^{K} \ell_{it} = \bar{\ell}, \qquad \sum_{i=1}^{K} s_i = \bar{s}. \tag{6}$$

Remark. In the baseline model, we separate R&D and production resources for expositional simplicity. As we show below, our results concerns the cross-sector allocation shares of R&D resources ($s_{it}/\bar{s}$), and our characterization is invariant to the level of R&D resources $\bar{s}$. Hence, all of our results hold in a richer model with factor mobility between R&D and production (see Online Appendix B.10).

2.2 Optimal Allocation of R&D Resources

In this section we characterize the optimal allocation of R&D resources in the economy. Consider a benevolent social planner who chooses the entire time path of worker and scientist allocations across sectors to maximize consumer utility. We can write the planner's problem as

$$V^{*}\left(\{q_{i0}\}\right) \equiv \max_{\{\ell_{it}, s_{it}\}} \int_{0}^{\infty} e^{-\rho t} \sum_{i=1}^{K} \beta_i \ln y_{it} \,\mathrm{d}t, \tag{7}$$

subject to the sectoral aggregator (3) for y_{it}, the flow of new innovation (4), law of motion for sectoral knowledge (5), and the resource constraints (6).

First, recognize from equation (3) that the planner's objective is log-linear in the allocation of production workers, implying the following lemma.

Lemma 1. *The planner allocates production workers in proportion to the consumption elasticity vector $\boldsymbol{\beta}$: for all t, $\ell_{it} = \beta_i \bar{\ell}$ for each sector i and variety ν.*

Lemma 1 simplifies the planner's problem into choosing how to allocate scientists only. Recall $\boldsymbol{\Omega} \equiv [\omega_{ij}]$ is the matrix that encodes the innovation network, and $\ln \boldsymbol{q}_t \equiv [\ln q_{it}]_{i=1}^{K}$ is the vector of log-knowledge stock at time t. Let $\gamma_{it} \equiv s_{it}/\bar{s}$ denote the share of scientists allocated to sector i at time t, and let $\boldsymbol{\gamma}_t$ denote the vector $[\gamma_{it}]_{i=1}^{K}$ that sums to one. Using equation (3) to express consumption in terms of production worker allocation and then applying Lemma 1, we rewrite the planner's problem in vector form as

$$\max_{\{\boldsymbol{\gamma}_t\} \text{ s.t. } \boldsymbol{\gamma}_t' \mathbf{1} = 1 \forall t} \psi \cdot \int_{0}^{\infty} e^{-\rho t} \boldsymbol{\beta}' \ln \boldsymbol{q}_t \,\mathrm{d}t \tag{8}$$

$$\text{s.t. } \mathrm{d} \ln \boldsymbol{q}_t \big/ \mathrm{d}t = \lambda \cdot \left(\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{\gamma}_t + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t\right). \tag{9}$$

We obtain (9) by substituting the innovation production function (4) into $\boldsymbol{q}_t$'s law of motion (5).

The planner's problem may seem intractable: the economy features a vector of state variables (sectoral knowledge stocks), and the law of motion involves dynamic network spillovers across sectors, meaning the allocation of R&D in any sector at any time affects the evolution of all state variables in all future times. Our formulation, however, is especially tractable: both the planner's objective function (8) and the law of motion (9) are log-linear in the state variables $\boldsymbol{q}_t$. Such tractability enables us to characterize the solution—the entire time path of optimal R&D allocation—in closed form. Later in Section 2.5 we generalize our analysis to a nonlinear setting.

Proposition 1. *Starting from any vector of initial knowledge stock $\boldsymbol{q}_0$, the optimal R&D allocation is time-invariant and follows, along the entire time path,*

$$\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}. \tag{10}$$

Proposition 1 shows the optimal cross-sector R&D allocation is time-invariant and follows $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$; the proportionality constant, $\frac{\rho}{\rho+\lambda}$, ensures that $\boldsymbol{\gamma}$ sums to one. To understand the intuition for the result, note that another way to write the optimal allocation vector of R&D resources $\boldsymbol{\gamma}'$ is:

$$\boldsymbol{\gamma}' \;\propto\; \boldsymbol{\beta}'\sum_{m=0}^{\infty}\left(\frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^m = \boldsymbol{\beta}'\left(\boldsymbol{I} + \frac{\boldsymbol{\Omega}}{1+\rho/\lambda} + \left(\frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^2 + \cdots\right).$$

That is, the Leontief inverse $\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$ in (10) can be written as a power series of $\frac{\boldsymbol{\Omega}}{1+\rho/\lambda}$. The first term in the infinite summation, $\boldsymbol{\beta}'\boldsymbol{I} = \boldsymbol{\beta}'$, captures how each sector's knowledge stock directly impacts consumer welfare through product quality. This term coincides with the optimal allocation of production workers (Lemma 1). The products between $\boldsymbol{\beta}'$ and subsequent terms in the power series capture the indirect effect of knowledge creation on consumer welfare, through future innovations and product quality improvements in network-connected sectors. Innovations in sector j benefit sector i by endogenously raising the efficiency of subsequent R&D in sector i, captured by the aggregator χ_{it} in equation (4) with elasticity ω_{ij}, which is the ij-th entry of the innovation network matrix $\boldsymbol{\Omega}$. Improved innovation efficiency in sector i further generates additional knock-on effects, as new knowledge in sector i facilitates future innovations in all sectors that benefit from sector i's knowledge stock; the higher-powered terms in the infinite summation capture these indirect effects.

Because network spillovers occur through sectoral knowledge stock, the flow of new knowledge through current R&D activities can only affect innovative efficiency and product quality in the future. Hence, the importance of network effects in the optimal R&D allocation is modulated by the discount rate ρ relative to λ, the sensitivity of knowledge growth to the flow of new innovation relative to the existing stock. The former (ρ) captures discounting of the future, and the latter (λ) captures how quickly those future benefits materialize. We refer to ρ/λ as the society's *effec-*

tive discount rate, which is a key parameter in determining optimal innovation allocation. When ρ/λ is high, the planner discounts future benefits heavily, and the network effects play a smaller role. In the limit as $\rho/\lambda \to \infty$, the planner becomes myopic, and the optimal R&D allocation is fully pinned down by the consumer preferences $\boldsymbol{\beta}$. Conversely, a more patient (low ρ/λ) planner allocates more R&D resources to sectors that benefit more sectors in the future, directly or indirectly. Proposition 1 implies that a patient planner directs R&D into basic science; an impatient planner directs R&D into consumer goods that may be peripheral in the innovation network, such as textiles and food products.[1]

2.3 R&D Allocation and Economic Growth

In this section we show how R&D allocation affects the economic growth rate along a balanced growth path (BGP). We demonstrate that the network's eigenvector centrality—what we call "innovation centrality"—is a sufficient statistic for evaluating the growth rate along a BGP and coincides with the growth-maximizing R&D allocation. We show the socially optimal R&D allocation $\boldsymbol{\gamma}$ converges to this growth-maximizing allocation when the planner is infinitely patient.

Definition 2. (Innovation Centrality) The vector of sectoral *innovation centrality*, $\boldsymbol{a} \equiv [a_i]_{i=1}^{K}$, is the dominant left-eigenvector of the innovation network $\boldsymbol{\Omega}$ with an associated eigenvalue of one, satisfying $\boldsymbol{a}' = \boldsymbol{a}'\boldsymbol{\Omega}$ and $\sum_{i=1}^{K} a_i = 1$.

Because $\boldsymbol{\Omega}$ is an irreducible, row-stochastic matrix, the innovation centrality vector $\boldsymbol{a}$ exists and is unique by the Perron-Frobenius theorem. We now show $\boldsymbol{a}$ is a key determinant of the BGP growth rate and coincides with the growth-maximizing R&D allocation.

Let $\boldsymbol{b}$ denote a generic vector of allocation shares with nonnegative entries and $\sum_{i=1}^{K} b_i = 1$.

Lemma 2. *Consider a BGP in which the aggregate consumption grows at a constant rate, with time-invariant allocations of production and R&D resources. Suppose R&D resources are allocated according to the vector $\boldsymbol{b}$ (i.e., $s_i/\bar{s} = b_i$); then, along the BGP, the growth rate of knowledge stock is the same across sectors and equals to*

$$g^q(\boldsymbol{b}) = \textit{constant} + \lambda \cdot \boldsymbol{a}' \ln \boldsymbol{b}, \tag{11}$$

where the right-hand side constant is $\lambda \cdot (\ln \bar{s} + \boldsymbol{a}' \ln \boldsymbol{\eta})$. The aggregate consumption growth rate is

$$g^y(\boldsymbol{b}) = \psi \cdot g^q(\boldsymbol{b}). \tag{12}$$

Lemma 2 analytically expresses the BGP growth rate of knowledge stock and the aggregate consumption as functions of the R&D allocation, $\boldsymbol{b}$. Along the BGP, the knowledge stock grows

[1] In Section B.1 of the Online Appendix, we provide an example with three sectors, and we analytically express the optimal allocation based on network structure and effective discount rate ρ/λ.

at the same rate $g^q(\boldsymbol{b})$ across sectors. The endogenous component of $g^q(\boldsymbol{b})$ is λ times the inner product between the innovation centrality ($\boldsymbol{a}$) and the vector of log-R&D allocation shares ($\ln \boldsymbol{b}$) (recall λ is the sensitivity of knowledge growth to the flow of new innovation relative to the existing stock). The exogenous component (the constant term) on the right-hand side of (11) shows that the growth rate is higher when R&D resources are more abundant (higher $\bar{s}$), when R&D leads to more new innovation flows (higher $\boldsymbol{\eta}$), and when a higher flow innovation (relative to the existing stock) leads to faster knowledge growth (higher λ). Since ψ parametrizes how knowledge translates into productivity, (c.f. equation 3), the growth rate of the aggregate consumption is simply ψ times the growth rate of knowledge stock across sectors.

Corollary 1. *The R&D allocation that maximizes the BGP consumption growth rate coincides with the innovation centrality* $\boldsymbol{a}$*, as it solves the following problem:* $\boldsymbol{a} = \arg\max_{\boldsymbol{b} \geq 0 \text{ s.t. } \mathbf{1}'\boldsymbol{b}=1} \boldsymbol{a}' \ln \boldsymbol{b}$.

This corollary highlights that innovation centrality $\boldsymbol{a}$ coincides with the growth-maximizing R&D allocation along a BGP. Intuitively, a_i captures the extent to which sector i's R&D activities contribute to economic growth, taking into account the network effects. Sectors with higher innovation centrality represent more fundamental technologies in the innovation network.

The corollary also demonstrates that the social planner does not necessarily choose the R&D allocation that maximizes growth. Unlike the socially optimal allocation $\boldsymbol{\gamma}$, which depends on the effective discount rate ρ/λ, the growth-maximizing allocation is equal to the innovation centrality and is independent of these parameters. Intuitively, the social planner maximizes the welfare of the consumer, who may prefer better quality products in the near future from consumption-intensive sectors (e.g., consumer goods such as textiles and food products), and knowledge in these sectors may not generate much knowledge spillovers for future innovations.

One can rewrite the optimal R&D allocation vector $\boldsymbol{\gamma}$ as the solution to the following fixed point equation, which demonstrates how $\boldsymbol{\gamma}$ varies with the effective discount rate ρ/λ:

$$\frac{\rho}{\lambda}\left(\boldsymbol{\gamma}' - \boldsymbol{\beta}'\right) + \boldsymbol{\gamma}'\left(\boldsymbol{I} - \boldsymbol{\Omega}\right) = \mathbf{0}'. \tag{13}$$

Equation (13) demonstrates that the optimal R&D allocation $\boldsymbol{\gamma}$ trades off between consumer preferences $\boldsymbol{\beta}$ and long-run growth captured by the innovation centrality $\boldsymbol{a}$, and ρ/λ modulates the relative importance of these two terms. When ρ/λ is large—an impatient planner—consumer preferences dominate ($\lim_{\rho/\lambda\to\infty} \boldsymbol{\gamma} = \boldsymbol{\beta}$). The planner places more weight on growth as ρ/λ declines; in the limit $\rho/\lambda \to 0$, equation (13) implies that $\boldsymbol{\gamma}'\left(\boldsymbol{I} - \boldsymbol{\Omega}\right) \to \mathbf{0}'$, thus the optimal allocation converges to the growth maximizing allocation ($\lim_{\rho/\lambda\to 0} \boldsymbol{\gamma} = \boldsymbol{a}$).

Proposition 2. *As the planner becomes infinitely patient ($\rho/\lambda \to 0$), the optimal R&D allocation converges to the innovation centrality, which is the growth-maximizing allocation:* $\lim_{\rho/\lambda\to 0} \boldsymbol{\gamma} = \boldsymbol{a}$. *As the planner becomes infinitely impatient ($\rho/\lambda \to \infty$), the optimal R&D allocation converges to the consumption elasticity vector:* $\lim_{\rho/\lambda\to\infty} \boldsymbol{\gamma} = \boldsymbol{\beta}$.

2.4 The Impact of R&D Allocation on Welfare

We now derive the welfare impact of R&D allocation, taking into account the transition dynamics.

Proposition 3. *For any initial knowledge stock* $\boldsymbol{q}_0$ *and any path of worker allocation* $\{\boldsymbol{\ell}_t\}$*, the difference in consumer welfare generated by two time-invariant R&D allocations* $\widetilde{\boldsymbol{b}}$ *and* $\boldsymbol{b}$ *is*

$$V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \widetilde{\boldsymbol{b}}\right) - V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \boldsymbol{b}\right) = \frac{\psi\lambda}{\rho^2}\boldsymbol{\gamma}'\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right). \tag{14}$$

Proposition 3 shows that the welfare difference resulting from two R&D allocation vectors can be expressed as the inner product between the optimal R&D allocation $\boldsymbol{\gamma}$ and the log-difference in R&D allocation vectors, multiplied by the scalar $\psi\lambda/\rho^2$. The result holds for any time path of worker allocation and any initial knowledge stock $\boldsymbol{q}_0$; hence, the Proposition can be used for welfare evaluation of policy counterfactuals that reallocate R&D resources across sectors.

Definition 3. (Consumption-Equivalent Welfare Gains from Adopting the Optimal R&D) Consider an economy with time-invariant R&D allocation $\boldsymbol{b}$ and the associated consumption path $\{y_t\}_{t\geq 0}$. The *consumption-equivalent welfare gains from adopting the optimal R&D allocation* $\boldsymbol{\gamma}$, is the scalar $\mathcal{L}(\boldsymbol{b})$ such that the consumer is indifferent between the consumption path $\{\mathcal{L}(\boldsymbol{b}) \times y_t\}_{t\geq 0}$ and the consumption path generated by reallocating R&D optimally according to the vector $\boldsymbol{\gamma}$, while holding worker allocation unchanged.

The scalar $\mathcal{L}(\boldsymbol{b})$ quantifies the welfare impact of reallocating R&D resources from $\boldsymbol{b}$ to $\boldsymbol{\gamma}$ across sectors. Note that, because the flow output is log-additive in the knowledge stock $\boldsymbol{q}_t$ and worker allocation $\boldsymbol{\ell}_t$ (equation 3), $\mathcal{L}(\boldsymbol{b})$ does not depend on the path of worker allocation. We next provide a simple formula for $\mathcal{L}(\boldsymbol{b})$.

Proposition 4. *The consumption-equivalent welfare impact of adopting the optimal R&D is*

$$\mathcal{L}(\boldsymbol{b}) = \exp\left(\frac{\psi\lambda}{\rho}\boldsymbol{\gamma}'\left(\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}\right)\right). \tag{15}$$

Proposition 4 shows that the consumption-equivalent welfare impact of reallocating R&D resources optimally, in logs, the inner product between the optimal allocation $\boldsymbol{\gamma}'$ and the log-difference between the optimal and the actual allocations, $(\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})$, multiplied by ψ, the elasticity of productivity to knowledge stock, and λ/ρ, the inverse of the effective discount rate. The inner product term, also known as the relative entropy, is a statistical distance measure of how $\boldsymbol{b}$ differs from $\boldsymbol{\gamma}$. Note that ρ/λ affects welfare through not only the proportionality constant but also the optimal allocation $\boldsymbol{\gamma}$ (as in Proposition 1).

2.5 General Functional Forms and Endogenous Innovation Network

The baseline model features an exogenous innovation network $\mathbf{\Omega}$, as the elasticities of each sector's innovation productivity to another sector's knowledge stock ($\omega_{ij} \equiv \frac{\partial \ln \chi_{it}}{\partial \ln q_{jt}}$) are exogenous structural parameters. The knowledge spillover dynamics (9) in the baseline model thus form a log-linear dynamical system. Log-linearity makes the baseline model especially tractable. Under this formulation, the optimal R&D allocation is time-invariant and holds along the entire transition path; the sufficient statistic for the welfare gains from changes in the R&D allocation also accounts for the gains along the entire transition path.

The log-linearity in the baseline model may appear restrictive at first, as it rules out the possibility that degree of knowledge spillovers depend endogenously on the levels of sectoral knowledge stock: as sector j accumulates knowledge q_{jt}, its contribution for sector i's innovation ($\frac{\partial \ln \chi_{it}}{\partial \ln q_{jt}}$) may rise or fall. Moreover, sector j's importance for consumption ($\frac{\partial \ln y_t}{\partial \ln y_{jt}}$) may also change endogenously. Such nonlinearity can be incorporated by modeling the consumption and innovation spillover elasticities ($\boldsymbol{\beta}$ and $\mathbf{\Omega}$) not as structural parameters but as objects that endogenously depend on the levels of knowledge stock. In such a richer, nonlinear environment, the optimal R&D allocation is no longer time-invariant; instead, it varies with the levels of knowledge stock across sectors (and depends also on the exogenous component η_i of each sector's innovation productivity).

We now show that our baseline results are actually more general. In nonlinear economic environments described above, our sufficient statistic serves as a first-order local approximation around a balanced growth path (BGP) for the potential welfare impact of reallocating R&D resources. The welfare impact arising from the endogenous changes in the consumption elasticities or the network structure (due to departure from log-linearity) is second-order in nature.

Specifically, consider replacing the log-linear consumption aggregator $y_t = \prod_{i=1}^{K} y_{it}^{\beta_i}$ in (2) with a general constant-return-to-scale function $y_t = \mathcal{Y}\left(\{y_{it}\}\right)$, and suppose the law of motion for sectoral knowledge stock follows

$$\dot{q}_{it}/q_{it} = \widetilde{f}\left(s_{it}\mathcal{X}_i\left(\{q_{jt}\}\right)\right) \tag{16}$$

where $\widetilde{f}\left(\cdot\right)$ is a concave function, and the R&D spillover function $\mathcal{X}_i\left(\cdot\right)$ satisfies homogeneous-of-degree-zero with positive cross-sector spillovers ($\partial \ln \mathcal{X}_i\left(\cdot\right)/\partial \ln q_{jt} > 0$ for $i \neq j$). We assume the spillovers are bounded, and without loss of generality we set the bound to one ($|\partial \ln \mathcal{X}_i\left(\cdot\right)/\partial \ln q_{jt}| \leq 1 \forall i, j$).

In this environment, consumer welfare given the path of R&D allocation $\{b_{it} \equiv s_{it}/\bar{s}\}$ is

$$\int_0^\infty e^{-\rho t} \ln \mathcal{Y}\left(\left\{q_{it}^{\psi}\ell_{it}\right\}\right) \,\mathrm{d}t, \quad \text{s.t. } \mathrm{d}\ln q_{it}/\,\mathrm{d}t = f\left(\ln\left(b_{it}\bar{s}\mathcal{X}_i\left(\{q_{jt}\}\right)\right)\right) \quad \forall i, \tag{17}$$

where $f\left(\cdot\right) \equiv \widetilde{f}\left(\exp\left(\cdot\right)\right)$, and ℓ_{it} is the mass of production workers allocated to sector i at time t.

In this environment, we can define the consumption and innovation spillover elasticities as

$$\beta_{it} \equiv \frac{\partial \ln \mathcal{Y}\left(\{y_{it}\}\right)}{\partial \ln y_{it}}, \qquad \omega_{ijt} \equiv \begin{cases} \frac{\partial \ln \chi_i(\{q_{kt}\})}{\partial \ln q_{jt}} & \text{if } i \neq j, \\ 1 + \frac{\partial \ln \chi_i(\{q_{jt}\})}{\partial \ln q_{it}} & \text{otherwise.} \end{cases}$$

These elasticities $\beta_{it}, \omega_{ijt} \in [0,1]$ are not structural parameters; instead, they are reduced-form objects evaluated at specific levels of allocation and knowledge stock. Precisely because these elasticities change endogenously, the optimal R&D allocation is no longer time-invariant and instead depends on the levels of knowledge stock across sectors.

Our results in the baseline model extends to this nonlinear environment locally around a BGP, where the allocations of worker and R&D resources are time invariant, and q_{it} grows at the same rate across all sectors. Let λ be the slope of $f(\cdot)$, which is common across all sectors. The consumption and innovation spillover elasticities $\boldsymbol{\beta} \equiv [\beta_i]$ and $\boldsymbol{\Omega} \equiv [\omega_{ij}]$ are time-invariant in a BGP; hence, we can define $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(I - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$ and scale its entries so that $\sum_i \gamma_i = 1$. In this nonlinear environment, the vector $\boldsymbol{\gamma}$ no longer represents the globally optimal R&D allocation (as in the baseline model); instead, it is proportional to the local elasticity of how R&D allocation affects welfare, as shown in the following result.

Proposition 5. *Consider an economy in a balanced growth path with R&D allocation $\boldsymbol{b}$. To first-order around the initial BGP, the consumption-equivalent welfare gain of moving from allocation $\boldsymbol{b}$ to $\widetilde{\boldsymbol{b}}$, accounting for transitional dynamics, is* $\exp\left(\frac{\psi\lambda}{\rho}\boldsymbol{\gamma}'\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right)\right)$.

The Proposition shows that the sufficient statistic for the welfare impact of R&D reallocation continues to hold as a first-order approximation around a BGP of the general nonlinear environment. Formally, as we show in Online Appendix A.7, the vector $\boldsymbol{\gamma}$ is proportional to the directional (Gateaux) derivative of welfare with respect to the R&D allocation. The welfare impact arising from the endogenous changes in the consumption elasticities or the network structure (due to departure from log-linearity) is second-order in nature.

2.6 Knowledge Spillovers from Abroad

We will later use our model to assess R&D allocations in real-world economies. As we show, some countries, like the U.S. and Japan, rely more on domestic knowledge spillovers and less on foreign knowledge spillovers, while other economies benefit more from foreign spillovers particularly from the technologically advanced ones. We now extend our model to incorporate this important cross-country heterogeneity, namely the reliance on knowledge spillovers from abroad, and examine how it affects the welfare impact of domestic R&D allocation.

2.6.1 Setup with Knowledge Spillovers from Abroad

We extend the closed economy model in Section 2.1 by introducing international knowledge spillovers and trade. The consumer now values both domestic and foreign goods with the constant-returns-to-scale preference aggregator $\mathcal{C}(\cdot)$:

$$V = \int_0^\infty e^{-\rho t} \ln \mathcal{C}\left(c_t^d, c_t^f\right) \, \mathrm{d}t,$$

where c_t^d is consumption of domestic goods and c_t^f is consumption of foreign goods. As in the closed economy model, domestic goods represent a Cobb-Douglas aggregation over sectoral goods (equations 2 and 3). The economy can import foreign goods c_t^f by exporting unconsumed domestic goods $\left(y_t - c_t^d\right)$. Let p_t^f be the relative prices of foreign goods. We impose trade balance:

$$p_t^f c_t^f = y_t - c_t^d. \tag{18}$$

Domestic innovation production benefits from foreign knowledge spillovers $\left\{q_{jt}^f\right\}_{j=1}^K$:

$$n_{it} = s_{it}\eta_i\chi_{it}, \qquad \text{where } \chi_{it} = \prod_{j=1}^K \left[(q_{jt})^{x_{ij}} \left(q_{jt}^f\right)^{1-x_{ij}}\right]^{\omega_{ij}}. \tag{19}$$

Like in the closed economy counterpart (4), n_{it} is the flow of new innovation generated by R&D resources s_{it} in sector i at time t; new innovation leads knowledge accumulation according to (5). χ_{it} is again the endogenous component of R&D efficiency in sector i. The difference here is that domestic R&D in sector i benefits from not only domestic knowledge q_{jt} in sector j but also foreign knowledge q_{jt}^f. The exponent x_{ij} captures the share of domestic contribution of knowledge spillover from sector j to sector i; when $x_{ij} = 1$ for all i, j, the innovation production function (19) coincides with the closed economy version (4). x_{ij}'s could differ across countries for various reasons. For instance, it could reflect a country's R&D comparative advantage; it could also reflect the degree to which foreign knowledge is accessible and may vary due to the cultural and political relations between the domestic economy and the rest of the world.

The domestic planner's problem is allocating workers and R&D resources to maximize domestic welfare, taking the time path of import prices $\left\{p_t^f\right\}$ and foreign knowledge $\left\{\boldsymbol{q}_t^f\right\}$ as given:

$$V^*\left(\boldsymbol{q}_0, \left\{\boldsymbol{q}_t^f, p_t^f\right\}_{t=0}^\infty\right) \equiv \max_{\{s_{it}, \ell_{it}\}} \int_0^\infty e^{-\rho t} \ln \mathcal{C}\left(c_t^d, c_t^f\right) \, \mathrm{d}t, \tag{20}$$

subject to the open economy innovation production function (19); trade balance (18); goods production functions (2) and (3); the law of motion for domestic knowledge (5); and resource constraints (6). To ensure the planning problem (20) is well-defined, we assume $\left|\frac{\mathrm{d} \ln q_{it}^f}{\mathrm{d}t}\right|$ is bounded, and $p_t^f > 0$ is bounded away from zero.

Remark. We make three remarks on the economic environment with foreign spillovers. First, while we assume x_{ij} to be exogenous and enters the innovation production function (19) log-linearly, our analysis holds as a first-order approximation in a richer environment where x_{ij} depends endogenously on the relative levels of domestic and foreign knowledge stock. This result, derived in Online Appendix B.7, is analogous to our closed-economy analysis in Section 2.5.

Second, we model international trade through the export of domestic bundle in exchange for the foreign bundle. This formulation significantly simplifies the exposition by removing the planner's incentive to use R&D allocation to manipulate the terms-of-trade (i.e., the relative prices of domestic products) and market power vis-a-vis foreign consumers.

Third, it is worth emphasizing that we do not analyze the optimal R&D allocation from the perspective of maximizing "global welfare". Doing so requires setting up a multi-country environment, which is certainly interesting and creates the opportunity to analyze a new set of other important issues—such as understanding cross-country R&D comparative advantages and strategic R&D policy—but it is beyond the scope of this paper.

2.6.2 Optimal R&D Allocation and Welfare with Knowledge Spillovers from Abroad

We now derive, in the setting with foreign spillovers, the optimal R&D allocation and the welfare impact of reallocating R&D resources.

Proposition 6. *Given paths of foreign knowledge and relative import prices* $\left\{\boldsymbol{q}_t^f, p_t^f\right\}_{t=0}^{\infty}$*, the solution to the open economy planner's problem (20) is time invariant and follows, along the entire time path,* $\ell_i/\bar{\ell} = \beta_i$ *and* $s_i/\bar{s} = \gamma_i$*, where*

$$\boldsymbol{\gamma}' = \xi^{-1}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^{-1}. \tag{21}$$

$\xi \equiv \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^{-1}\mathbf{1}$ *is a constant that ensures* $\sum_i \gamma_i = 1$*;* $\boldsymbol{X} \equiv [x_{ij}]$ *is the matrix encoding the share of domestic contribution to cross-sector knowledge spillovers;* $\circ$ *is the Hadamard product.*

Proposition 6 generalizes Proposition 1 in Section 2.2 to an open economy. The ij-th entry of the Leontief inverse $\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^{-1} \equiv \sum_{m=0}^{\infty}\left(\frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^m$ summarizes the network spillover effects from additional *domestic* knowledge in sector j on subsequent *domestic* innovation in sector i. Unlike in the closed economy, each round of network effect is no longer captured by the innovation network $\boldsymbol{\Omega}$ but is instead captured by $\boldsymbol{\Omega}\circ\boldsymbol{X}$: in the presence of knowledge spillovers from abroad, domestic R&D only contributes partially to the total knowledge spillovers from sector j to sector i; the elasticity of innovation efficiency in sector i with respect to domestic knowledge in sector j is captured by the ij-th entry of $\boldsymbol{\Omega}\circ\boldsymbol{X}$ (i.e., $\frac{\partial \ln \chi_i}{\partial \ln q_j} = \omega_{ij}x_{ij}$).

Proposition 6 highlights that countries with more self-contained innovation networks—such the U.S. and Japan, as we show later, where R&D builds more heavily on domestic rather than on foreign knowledge—should allocate more R&D to network-central sectors. Conversely, countries that benefit more from foreign spillovers should direct R&D toward sectors that account for greater domestic value-added. Using our intuition from the closed economy Proposition 1, it is as if economies with self-contained innovation networks have patient planners while economies reliant on foreign knowledge have impatient planners. To see this, consider an economy in which the domestic share of knowledge spillovers is constant across all sector-pairs, $x_{ij} = x$. The Leontief inverse in (21) simplifies to $\left(\boldsymbol{I} - x \cdot \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$. Greater reliance on foreign knowledge (lower x) is therefore isomorphic to a higher discount rate ρ in a closed-economy, as if the planner place less value on long-term innovation spillovers.

The proportionality constant ξ in equation (21) ensures γ_i sums to one. It is a measure of R&D self-sufficiency. $\xi \leq 1$ in open economies and is decreasing in foreign dependence; $\xi = 1$ only if the economy does not benefit from any foreign spillovers ($x_{ij} = 1$ for all i, j).

Our next result provides the consumption-equivalent welfare impact of adopting the optimal R&D allocation, extending our closed-economy result in Proposition 4.

Proposition 7. *Consider an open economy with R&D self-sufficiency measure ξ and given paths of foreign knowledge and relative import prices $\left\{\boldsymbol{q}_t^f, p_t^f\right\}_{t=0}^{\infty}$. For any path of worker $\{\boldsymbol{\ell}_t\}$ and time-invariant R&D allocation $\boldsymbol{b}$, the consumption-equivalent welfare impact of adopting the optimal R&D allocation is*

$$\mathcal{L}(\boldsymbol{b}; \xi) = \exp\left(\psi \xi \frac{\lambda}{\rho} \boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})\right). \tag{22}$$

Relative to the closed-economy counterpart (15), the open-economy formula (22) depends additionally on ξ, the R&D self-sufficiency measure defined in Proposition 6. This term formalizes the notion that, in economies where domestic R&D matters less for long-run spillovers (lower ξ), suboptimally allocating domestic R&D is also less consequential for welfare.

2.7 Additional Results and Theoretical Extensions

We now discuss several additional theoretical results and extensions. Section 2.7.1 incorporates a production network into the baseline model. Section 2.7.2 embeds an innovation network into a semi-endogenous growth setting. Section 2.7.3 discusses potential inefficiencies in a decentralized equilibrium. Section 2.7.4 briefly describes several other extensions in the Online Appendix.

2.7.1 Production Network

Our baseline model features a simple production structure where all goods are produced directly from labor. We now discuss how to tractably incorporate a canonical production network into our framework. More details of this extension are provided in Section B.2 of the Online Appendix.

Conceptually, for the optimal R&D allocation $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$, the presence of a production network requires a different construction for the $\boldsymbol{\beta}$ vector, but the innovation network $\boldsymbol{\Omega}$

term is unaffected. Formally, the $\boldsymbol{\beta}$ vector should capture the elasticity of aggregate consumption with respect to the productivity in each sector ($\frac{\partial \ln y_t}{\partial \ln q_{it}^{\psi}}$); in the presence of a production network, it should reflect not only the consumer preferences but also the production network structure. With this adjustment, our main results continue to hold in this environment.

Specifically, suppose the production of good i requires other goods as intermediate inputs, as in the canonical production network model (Acemoglu et al., 2012):

$$y_{it} = \left(q_{it}^{\psi}\ell_{it}\right)^{\alpha_i} \prod_{j=1}^{K} m_{ijt}^{\sigma_{ij}}, \quad \alpha_i + \sum_{j=1}^{K}\sigma_{ij} = 1, \tag{23}$$

where m_{ijt} is the quantity of good j used for the production of good i, α_i is sector i's output elasticity to value-added, and σ_{ij} is sector i's output elasticity to input j. The baseline model is a special case with $\sigma_{ij} = 0$ for all i, j.

Unlike in our baseline model, the elasticity of the aggregate consumption with respect to sectoral productivity is no longer the consumption elasticity β_i; instead, standard results from the production network literature (e.g., see Hulten, 1978) imply that $\frac{\partial \ln y_t}{\partial \ln q_{it}^{\psi}} = \hat{\beta}_i \equiv \alpha_i \delta_i$, where $\boldsymbol{\delta}' = \boldsymbol{\beta}' \left(\boldsymbol{I} - \boldsymbol{\Sigma}\right)^{-1}$ is the influence vector, and $\boldsymbol{\Sigma} \equiv [\sigma_{ij}]$ is the input-output elasticity matrix. If the marginal product of labor is equalized across sectors, then $\hat{\beta}_i$ is the share of labor allocated to sector i. Our main results from previous sections, including the optimal R&D allocation and the welfare sufficient statistic, continue to hold when we use the vector $\hat{\boldsymbol{\beta}}$ in place of $\boldsymbol{\beta}$. In our empirical exercises, we measure this vector using sectoral value-added as a share of GDP, thereby accounting for the production network.

2.7.2 Semi-Endogenous Growth

Our baseline model features endogenous growth: a positive growth rate of aggregate output along a balanced growth path in the absence of population growth. This is because the R&D technology features aggregate constant-returns-to-scale in sectoral knowledge stock. To embed our innovation network formulation into a semi-endogenous growth setting with a constant growth rate in the total measure of scientists $\bar{s}_t = \bar{s}_0 e^{\bar{g}t}$, consider replacing the knowledge stock evolution equation (5) with

$$\dot{q}_{it}/q_{it} = \lambda \ln\left(n_{it}/q_{it}^{1+\kappa}\right),$$

where $\kappa \geq 0$ captures the rate at which proportional improvements in knowledge are getting harder to find (Bloom et al. 2020, Jones 2022). The knowledge law of motion (9) becomes

$$\mathrm{d}\ln \boldsymbol{q}_t / \,\mathrm{d}t = \lambda \cdot \left(\ln \boldsymbol{\eta} + \ln \bar{s}_0 + \bar{g}t + \ln \boldsymbol{\gamma}_t + \left(\boldsymbol{\Omega} - (1+\kappa)\,\boldsymbol{I}\right)\ln \boldsymbol{q}_t\right).$$

Our baseline model corresponds to the special case where $\kappa = \bar{g} = 0$. When $\kappa > 0$, the R&D technology features aggregate decreasing-returns-to-scale, and the long-run growth rate of the economy is fully determined by the growth rate $\bar{g}$ of scientists. In Online Appendix B.3, we show that the optimal R&D allocation follows $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}$, and the consumption-equivalent welfare impact of adopting the optimal allocation is $\mathcal{L}(\boldsymbol{b}) = \exp\left(\frac{\psi\lambda}{\rho+\kappa\lambda}\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b})\right)$. Compared to the formulas under endogenous growth, the semi-endogenous-growth version replaces the discount rate ρ with $\rho + \kappa\lambda$. Intuitively, in the presence of aggregate decreasing returns, growth in knowledge stock raises the difficulty of future R&D; from a welfare perspective, it is *as if* future growth are discounted at a higher rate.

2.7.3 Potential Inefficiency in A Decentralized Market

Why may a decentralized market not allocate R&D resources efficiently? In an innovation network, knowledge is a public good, as knowledge creation benefits subsequent R&D in other sectors and all future periods. To the extent that innovators do not fully internalize the future benefits,[2] a decentralized market does not implement the efficient R&D allocation. To demonstrate the potential inefficiency, Online Appendix B.4 constructs a decentralized equilibrium in which innovators conduct R&D only in pursuit of profits—each innovation is patented, thereby granting the innovator a temporary stream of profits until replaced by a future innovation—disregarding any beneficial spillovers their R&D activities may provide in the future. As we show, the decentralized allocation of R&D resources along a BGP follows the consumption elasticity $\boldsymbol{\beta}$, which can be efficient only if the society is completely myopic ($\rho/\lambda \to \infty$).

This illustrative decentralized equilibrium lacks many real-world features of the market for innovation (e.g., multi-sector firms, mergers and acquisitions, and patent licensing). Nevertheless, it is important to note that, by comparing the R&D allocation in the data to the first-best, our notion of allocative efficiency—measured by the consumption-equivalent welfare impact of reallocating R&D optimally—does not require that we take a stance on firms' equilibrium conduct; instead, it directly calculates the welfare impact of reallocating R&D based on the economic environment specified in Section 2.1.

2.7.4 Other Theoretical Extensions

The Online Appendix includes a number of additional extensions that generalize our main results to various economic environments.

Section B.5 shows that our results extend naturally to a setting where the planner is constrained and can reallocate resources across only a subset of sectors: the constrained-optimal resource allocation within the subset is proportional to the unconstrained optimal allocation $\boldsymbol{\gamma}$.

The open economy analysis in Section 2.6 considers a domestic planner who takes the paths

[2] It is worth noting that patents do not necessarily correct for knowledge externalities: while the patent holder has the exclusive right to use the invention, it does not preclude others from benefiting from knowledge spillovers.

of foreign knowledge as given. Section B.6 considers a domestic planner who internalizes the impact of domestic allocations on foreign variables. Section B.7 concerns a setting where both the innovation network ($\mathbf{\Omega}$) and the elasticities of foreign spillovers ($\boldsymbol{X}$) are endogenous.

Section B.8 considers an environment with sector-specific λ_i, the sensitivity of knowledge growth to the flow of new innovation relative to the existing stock.

Section B.9 considers a setting where the row-sums of $\mathbf{\Omega}$ may differ (and not equal to one). We show that the knowledge stock in each sector no longer grows at the same rate along a BGP; instead, the vector of growth rates of q_{it} form the right-Perron vector of $\mathbf{\Omega}$.

Section B.10 allows for factor mobility between production and R&D.

3 Data

This section describes the data for our empirical analyses. We use patent citation data across sectors and countries to construct the global innovation network. We also use data on sectoral production, final use, and R&D. Here we briefly describe how we construct and harmonize these data. Section C of the Online Appendix provides more details.

3.1 Data on Patents

U.S. Patents U.S. patent data are from the United States Patent and Trademark Office (USPTO).[3] This database provides detailed patent-level records for nearly seven million patents granted by the USPTO since 1976. The data include, for each patent, the application and grant years, the technology classifications based on the International Patent Classification (IPC) system, and the geographic locations of the patent assignee and patent inventors (the former holds legal ownership rights to the patent while the latter may not). Central to our network analysis, we observe each patent's citations of prior patents as well as the citations it receives from subsequent patents up to 2020, the year we extracted the data. In our empirical analysis, we use patents filed before the end of 2014 to mitigate the right-truncation problem, since patents filed more recently may still be in the approval process.

Global Patents To capture global innovation, we use Google Patents' global patent data, which contain information on more than 36 million patents from over 40 major patent authorities around the world, including those in the U.S., the European Union, Japan, and China, among others, starting from the 18th century (data prior to the 1970s have limited coverage). Google Patents global innovation data are constructed based on the raw data records at DOCDB (EPO worldwide

[3] We accessed the patent data via the USPTO PatentsView platform at https://www.patentsview.org/download/.

bibliographic data), which are the same records underlying other global patent datasets such as the PATSTAT database. As a result, Google Patents's data coverage and variable quality are nearly identical to those of PATSTAT. We choose to use Google Patents data for our main analysis because they are free of charge for any researcher, and we compare Google Patents to PATSTAT in detail in the Online Appendix D.

For each patent, Google Patents supplies information similar to the USPTO data described above. We assign each patent to a geographical unit using the country of residence of its inventors(s), country of residence of the assignee(s), and country of the patent authority, in that order. When a patent is associated with inventors or assignees from multiple countries, we attribute the patent to these countries assuming fractional and equal weight per assignee/inventor.

A major challenge when working with international patent data is multi-filing: to protect intellectual properties, it is common practice for innovators to file the same invention with multiple patent authorities in different countries, forming what is called a "patent family." To avoid double counting these inventions, our analysis uses only the first patent filed in each family when counting new innovation, while attributing all citations made to a whole family to this first patent. To identify patent family, we use the patent family ID assigned by Google Patents, self-reported multi-filing status, and the unique identifier for patents filed under the Patent Cooperation Treaty, which is an international law treaty aimed at protecting innovations across countries.

We measure the number of patents produced in a country-sector-year, both the raw counts and with quality adjustments using the number of citations each patent received. To capture actual patent timing, we use the year a patent was filed rather than granted.

3.2 Data on Production, Final Use, and R&D Allocation

Production and Final Use In our cross-country analysis, for each country and sector, we use the World Input-Output Database (WIOD, Timmer et al., 2015) to extract sectoral information. The data cover the years 2000–2014 and 43 major economies, which altogether represent more than 85% of world GDP. WIOD's sectoral categorization follows the two-digit International Standard Classification (ISIC) revision 4, with a total of 56 sectors covering the entire production spectrum, including primary, manufacturing, and service sectors. We obtain data on value-added (gross value-added, "VA"), employment ("empe"), output (gross output, "GO"). We also obtain information on intermediate inputs, value used for consumption, imports, and exports. For the U.S., we also obtain more detailed sectoral production, consumption, and import-export data, comprising 181 sectors from 1990 to 2019, from the Bureau of Labor Statistics (BLS).

Sectoral R&D Allocation Our quantitative analysis uses data on R&D allocation across different technology classes in each country. There is no standard database to exhaustively measure such information. Our primary measure relies on aggregating firm-level R&D expenditures to the country-sector-year level, based on three widely used firm-level data sets: Compustat, Worldscope, and Datastream. Combined, these data cover more than 110,000 global firms located in 160 countries and account for over 95% of the world's total market capitalization. For multinationals, we first attribute the firm-level R&D expenditures to IPC-country level in proportion to each firm's shares of patents in each IPC-country, following Griffith, Harrison, and Van Reenen (2006), and then aggregate to IPC-country-year level.

Our primary measure of sectoral R&D is imperfect, as the firm-level data sets oversample large firms and have potentially different reporting standard across countries; we also miss R&D inputs from public sectors. Nevertheless, it is important to note that, as our theory concerns the cross-sector R&D allocation, what matters for our quantitative analysis later is the allocation shares of R&D resources across sectors in each country and not the aggregate R&D levels; any mismeasurement that affects all sectors proportionally should have no quantitative impacts. As robustness checks, we later show that our primary measure of R&D allocation shares correlates strongly with two independent sources of R&D data, thereby giving us confidence in using our measure for quantitative analysis. The first robustness check calculates cross-sector R&D allocation using the innovation output (which is better measured) rather than input: the number of patents produced in each country-IPC (or country-sector) divided by total number of patents produced in that specific country (correlation with our primary measure of sectoral R&D averages to 0.74 across countries; see Appendix Table A.1). The second robustness check utilizes the OECD Analytical Business Enterprise Research and Development (ANBERD) Database (Machin and Van Reenen, 1998), which has country-sector-level R&D information. Relative to our primary R&D measure, the ANBERD Database has more limited country-year coverage and relies more on imputations from firm-level surveys. Our primary R&D measure also allows us to explicitly and transparently attribute R&D of multi-sector or multinational enterprises to different sectors and countries. Nevertheless, for all the major economies in both data sets, R&D allocation from ANBERD is highly correlated with our primary measure (see Appendix Table A.1), and our quantitative results are consistent using both data.

3.3 Concordance

Patent data are classified according to the IPC system, which is distinct from the classifications in our sectoral data. We build concordance between these two data types to construct sectoral measures of innovation activities and reversely to project sectoral measures into technology class levels. To project patents from IPC onto sectors, we leverage the sectoral classifications covered in the three firm-level data sets described above. For the U.S., we link the USPTO patent database to Compustat using the bridge file provided by Kogan et al. (2017) and Ma (2021). Firms are classified by the North American Industry Classification System (NAICS) codes, which are then mapped to BLS sectors using the crosswalk file provided by the BLS website.[4] For the global analysis, we follow the analogous procedure and match Google Patents with global firm data from the Worldscope and Datastream databases. This process provides each patenting firm's International Standard Industrial Classification (ISIC), which can then be accurately mapped to WIOD that is also organized using the ISIC system. To reverse-project country-sector-year level measures onto country-IPC-year, we use the sector-IPC mapping provided in Lybbert and Zolas (2014).

We provide details on these matching procedures and the representativeness of using innovation measures aggregated from firm-level data in Section C of our Online Appendix.

4 Innovation Network and Knowledge Spillovers

In this section, we build several key data elements that will be used in our main quantitative analysis in Section 5. We first construct the innovation network $\mathbf{\Omega}$ and discuss its empirical properties. We then empirically validate a key mechanism in our model, that knowledge spillovers occur through innovation networks both domestically in the U.S. and globally.

4.1 Innovation Network

Constructing the Innovation Network We construct the innovation network from patent citations. First, we build the sector-to-sector innovation network. Let $Cites_{ijt}$ denote the total number of times that patents in sector i cite patents in sector j, among all patents filed in year t. As a baseline construction, we follow (Acemoglu et al., 2016) and define ω_{ijt} as the share of total citations patents in sector i made to sector j in year t:

$$\omega_{ijt} \equiv \frac{Cites_{ijt}}{\sum_{k=1}^{K} Cites_{ikt}}. \tag{24}$$

The object ω_{ijt} measures the extent to which upstream sector j's prior knowledge benefits innovation in sector i. The matrix $\mathbf{\Omega}_t$, whose ij-th entry is ω_{ijt}, captures the knowledge flow network we refer to as the *innovation network*. In a global setting, all subscripts will include additional country dimensions indicating the countries of sectors i and j, and the innovation network will measure the extent to which upstream country's sector j's prior knowledge benefits innovation in focal country's sector i. We can construct a country-specific network using patents from each country; we can also include patents from a time window broader than one year, such as using all patents over five or ten years.

The sector-to-sector innovation networks appear to be stable over time and across countries,

[4] The crosswalk can be accessed at: https://www.bls.gov/emp/documentation/crosswalks.htm.

suggesting their ability to capture some deep relations between sectoral innovation and technology diffusion. Table A.5 of the Online Appendix shows the serial correlation of entries in $\mathbf{\Omega}_t$ is near perfect a decade apart and remains above 0.8 even three decades apart. Table A.6 of the Online Appendix demonstrates that the innovation network constructed by pooling patents from all countries near perfectly correlates with the U.S.-specific network (correlation 0.97) and highly correlates (correlation ≈ 0.8) with country-specific innovation networks from Japan, China, Germany, Canada, the United Kingdom, and France. These findings imply that decisions about country and time specificities of the innovation network do not materially affect our analysis.

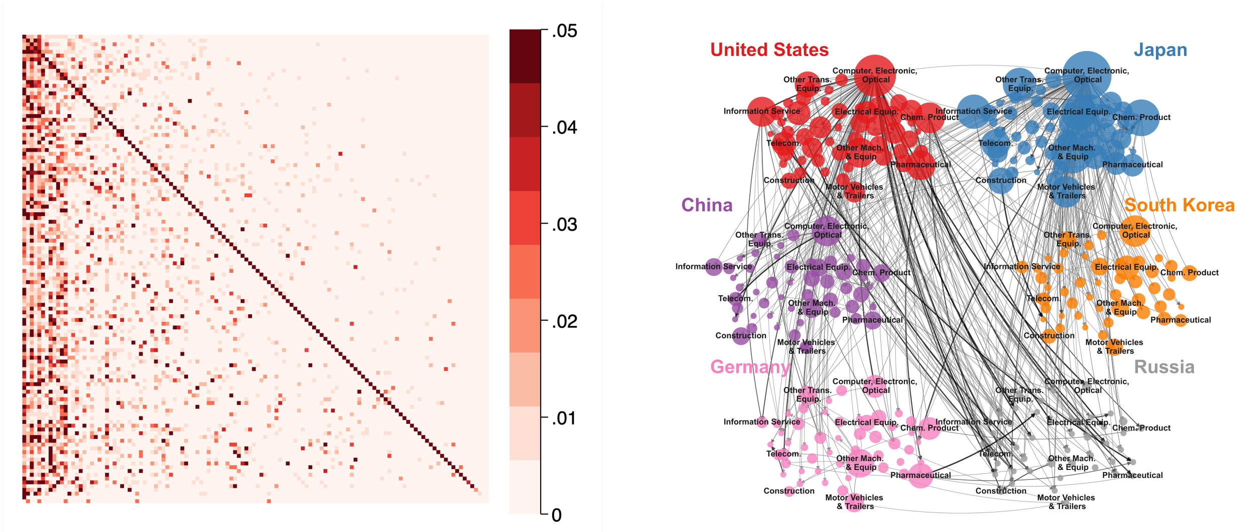
We should point out that the knowledge spillover network is inherently difficult to measure. Despite the evidence (Section 4.2 below in particular) for the usefulness of our citation-based construction, it is not perfect. In later analysis, we provide several alternative constructions of the network, such as weighting connections by the quality (total citations) of the cited or citing patents, to focus only on spillovers among major patents; we also create an innovation network $\mathbf{\Omega}$ with entries proportional to citations, without the normalization by each row's sum. Our quantitative findings are robust across all of these specifications.

Visualizing the Innovation Network Figure 1, Panel (a) visualizes the innovation network by plotting the matrix $\mathbf{\Omega}$ of 2010 as a heatmap. Each row and each column is a 3-digit IPC class, where the color in the i-th row and j-th column corresponds to ω_{ij} using the colormap listed to the right of the figure. Sectors are sorted by decreasing innovation centrality, the empirical properties of which we will formally discuss below. A key feature is that IPC classes follow a hierarchical structure: the innovation network is highly asymmetric, and there is a "pecking order" across sectors. Innovation-central sectors account for a disproportionate share of citations from all other sectors (columns are dense on the left but become progressively sparser to the right), yet these innovation-central sectors do not significantly cite noncentral sectors (rows are sparse on the top but become progressively denser toward the bottom).

Figure 1, Panel (b) visualizes the global innovation network by plotting each country-sector as a node, with size drawn in proportion to the total patent counts in our sample. An arrow from country m sector j to country n sector i indicates knowledge flow from mj to ni, with arrow width drawn in proportion to the share of ni's citations to mj. For visual clarity, only the largest countries and sectors are shown. Several patterns emerge from this figure. First, Japan and the U.S. produce the most patents in our sample. Second, the U.S. receives significantly more foreign citations than any other economy in our sample; it is a major knowledge exporter and only a minor knowledge importer.

Figure 1. Visualizing the Innovation Network

(a) IPC-to-IPC (131×131) network **Ω** **(b)** The global innovation network



Notes. The left panel visualizes the IPC-to-IPC (3-digit level) network **Ω** as a heatmap, with darker colors representing larger matrix entries; sectors are ordered according to their innovation centrality. The right panel visualizes the global innovation network for six economies with the highest total patent output in our sample. Each node is a country-sector, with size drawn in proportion to patent output. Arrows represent knowledge flows, with width drawn in proportion to citation shares.

The Innovation Network Weakly Correlates with Input-Output Networks The innovation network **Ω** encodes cross-sector linkages via knowledge spillovers. Another prominent type of cross-sector linkages occur through input-output relations, as sectors purchase intermediate inputs from one another during production. Table 1 shows that innovation and production networks are only weakly correlated. In other words, the two network relations capture different connections across sectors. Specifically, for each of the top ten countries ranked by total patent output, we compute the industry-by-industry input-output expenditure share matrix, which is a row stochastic matrix (as is **Ω**) commonly used to represent input-output relationships. Table 1 presents the correlation between entries in **Ω** and those in the input-output matrix. The correlation is weak (<0.35) in all economies.

Table 1. Correlations Between Country-Level Innovation Network and Production Network

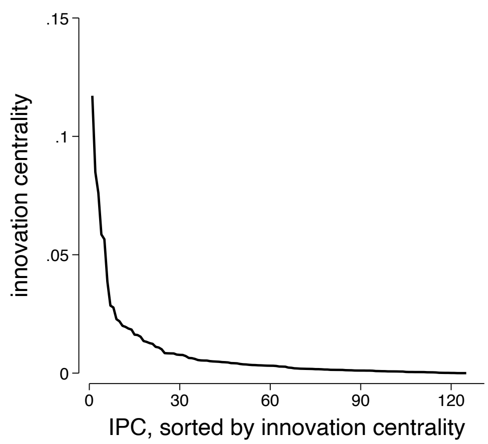
US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
0.32	0.28	0.35	0.31	0.23	0.19	0.36	0.41	0.29	0.22

Notes. This table presents the correlations between the country-level innovation network matrix and the country-level input-output expenditure share matrix for the top 10 countries ranked by total patent counts during 2010–2014.

Innovation Centrality Across Sectors We provide some descriptive statistics of the innovation centrality $\boldsymbol{a}$, which is the dominant left eigenvector of the innovation network $\boldsymbol{\Omega}$. Recall that in our model, $\boldsymbol{a}$ is also the R&D allocation vector that maximizes the growth rate of a closed economy (Corollary 1) and is an important determinant of the optimal R&D allocation. The left panel of Figure 2 plots the innovation centrality a_i across 3-digit IPC sectors using the 2010 U.S. innovation network, where sectors are ordered along the x-axis in descending a_i. The figure shows innovation centrality is highly heterogeneous across sectors. To maximize economic growth, the most innovation-central sector should be allocated about twice as many R&D resources as the 5th sector ranked by a_i, about ten times as many as the 20th sector, and about 30 times as many as the 50th sector. The right panel of Figure 2 identifies the top 10 IPC classes; these include several technological classes related to medical science, computing, semiconductors, and electric communication technologies, among others.

Figure 2. Innovation Centrality and Key Sectors

(a) Innovation Centrality Across IPCs



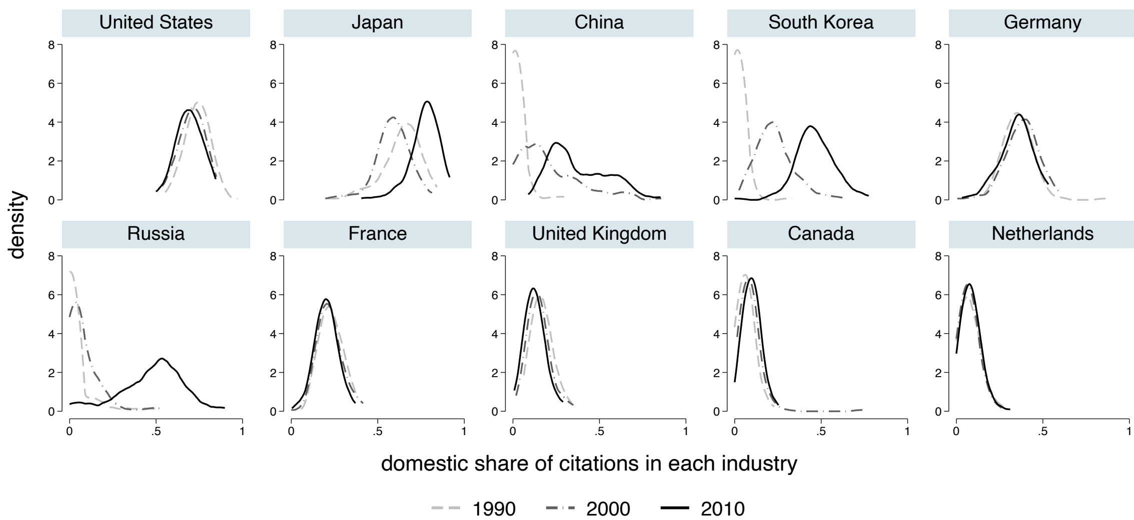
(b) Top 10 IPCs By Innovation Centrality a_i

1	A61	medical or veterinary science; hygiene
2	G06	computing; calculating or counting
3	H01	basic electric elements
4	H04	electric communication technique
5	G01	measuring; testing
6	B60	vehicles in general
7	G02	optics
8	B01	physical or chemical processes or apparatus in general
9	C08	organic macromolecular compounds; their preparation or chemical working-up; compositions based thereon
10	F16	engineering elements or units; general measures for producing and maintaining effective functioning of machines or installations; thermal insulation in general

Notes. This figure presents the innovation centrality of different technology classes categorized using IPC. Panel (a) plots innovation centrality a_i across 3-digit IPC sectors ranked in descending order based on a_i. Panel (b) lists the top 10 IPC classes by their innovation centrality.

Cross-Country Linkages in the Innovation Network How much do countries benefit from foreign knowledge? To answer this, for each country m, sector i, and year t, we compute the domestic share of citations made by patents in country-sector mi. Figure 3 shows the distribution of domestic citation shares across all sectors for the ten economies with the highest patent counts in our sample, for the years 1990, 2000, and 2010. The U.S. relies relatively sparingly on foreign knowledge: consistently across these three decades, about 70% of citations by U.S. patents are made to other U.S. patents. In contrast, citations made to foreign patents account for the vast majority of citations by all other economies except Japan, suggesting these economies benefit significantly from foreign knowledge, most notably from the U.S. The Japanese self-citation shares increased over time on average, from 65% in 2000 to 77% in 2010. Declining foreign reliance over time is also observed for China and South Korea, although their levels of foreign reliance remain high.

Figure 3. Cross-Sector Distribution of Domestic Citation Shares by Country



Notes. This figure presents the distribution of each country's domestic citation shares across sectors, showing the distribution using 1990, 2000, and 2010 data. Sector definitions follow 3-digit IPC classes. Domestic citation share for each country-sector is defined as the number of citations made to domestic patents as a share of total citations made by new patents invented in that each country-sector.

4.2 Knowledge Spillovers Through the Innovation Network

We now provide evidence that knowledge spillovers occur through the innovation network, with the purpose of validating our model mechanism and the innovation network construction. We build on Acemoglu, Akcigit, and Kerr (2016)—which provide similar evidence for the U.S. domestic network—and extend using instrumental variables (IVs) and to the global setting. The IVs are constructed based on time-varying sectoral exposure to tax-induced user cost of R&D (Wilson, 2009 and Thomson, 2017); they help isolate movements in patent output driven by knowledge spillovers and not by common shocks to connected sectors (Manski, 1993 and Bloom, Schankerman, and Van Reenen, 2013). We find evidence for directional knowledge spillovers: each sector's innovation output responds only to past upstream innovations and does not respond to past innovation from downstream sectors even though they are also connected. We also show that relative to input-output linkages, the innovation network is a significantly stronger channel through which knowledge spillovers take place.

4.2.1 U.S. Evidence

We first test the mechanism in the U.S. using an empirical specification derived by our model, treating the U.S. as a closed economy. Specifically, integrating our law of motion (5) over time, we can express the knowledge stock as $\ln q_{jt} = \int_0^\infty e^{-\lambda s} \ln n_{j,t-s}\,\mathrm{d}s$. The innovation production function (4) further implies a log-linear relationship among sector i's new patents, sectoral R&D, and past patents from other upstream sectors:

$$\ln n_{it} = \ln \eta_i + \ln s_{it} + \underbrace{\lambda \sum_{j=1}^{K} \omega_{ij} \left(\int_0^\infty e^{-\lambda s} \ln n_{j,t-s}\,\mathrm{d}s \right)}_{\substack{\equiv \ln \chi_{it}\text{, the aggregation of knowledge that} \\ \text{benefits innovation in sector } i \text{ at time } t}}. \tag{25}$$

Equation (25) implies that, after controlling for sectoral R&D expenditures, past patents $\ln n_{j,t-s}$ in sector j influence new patent output in sector i through the innovation network ω_{ij}. Equation (25) also shows that, importantly, knowledge spillovers' effect is directional: the knowledge flow from sector j to sector i operates through ω_{ij} and not ω_{ji}.

We test the discrete-time analogue of (25) by constructing the knowledge aggregator χ_{it} from past patents. As a baseline measure, for each sector i, we enumerate over all sectors j from which knowledge flows to i, aggregating j's log patent counts $\ln n_{j,t-\tau}$ in the past 10 years ($1 \leq \tau \leq 10$), weighted by $\omega_{ij,t-\tau}$, the share of citation made from i to j in the corresponding year:

$$\text{Knowledge}_{it}^{Up} \equiv \textstyle\sum_{j \neq i} \sum_{\tau=1}^{10} \omega_{ij,t-\tau} \ln n_{jt-\tau}. \tag{26}$$

$\text{Knowledge}_{it}^{Up}$ captures the stock of past knowledge "upstream" of sector i and is the empirical counterpart to $\ln \chi_{it}$. We then perform the following regression, with sector and year fixed effects:

$$\ln n_{it} = \beta_1 \times \text{Knowledge}_{it}^{Up} + \beta_2 \times \ln R\&D_{i,t-1} + \xi_i + \xi_t + \epsilon_{it}, \tag{27}$$

where n_{it} is the number of patents filed in sector i year t that are eventually granted. $R\&D_{i,t-1}$ is the R&D stock, which is the accumulated R&D expenses over the past five years using a decay rate of 15%, following Hall et al. (2005) and Bloom et al. (2013). These results are robust to using alternative measures of R&D such as concurrent or lagged R&D expenditures, and R&D stocks calculated using 5% or 10% knowledge decay rate.

Note that, when constructing the upstream knowledge aggregator (26), we exclude the lagged patent output from each sector itself. Doing so ensures the coefficient β_1 in regression (27) is not driven by serially correlated shocks to sectoral patent output, but our results are robust to including lagged patent output from own sector as an additional regressor (Appendix Table A.8). Also note that, theoretically, the knowledge aggregator in (25) discounts past patents exponentially. Our empirical construction (26) features a discrete cutoff window ($\tau \leq 10$ years) to be agnostic about the discount factor λ, but our results are robust to alternative values of τ (see Appendix Table A.9 for $\tau = 5$ and 20), exponential discounting with an annual discount rate of 15% (Appendix Table A.10), or estimating the impact of past upstream patents nonparametrically at different time lags (Appendix Figure A.9).

Table 2. U.S. Evidence of Innovation Spillovers Through the Innovation Network

$Y=$	ln(Patents)				ln(Cites)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Knowledge_{it}^{Up}$	0.555***	0.605***	0.509***	0.583**	0.790***	0.840***	0.756***	0.917***
	(0.174)	(0.194)	(0.169)	(0.269)	(0.197)	(0.207)	(0.192)	(0.289)
$ln(R\&D\ Stock)_{i,t-1}$	0.426***	0.433***	0.410***	0.408***	0.340***	0.347***	0.328***	0.206
	(0.100)	(0.101)	(0.096)	(0.111)	(0.114)	(0.114)	(0.111)	(0.133)
$Knowledge_{it}^{Down}$		-0.112				-0.110		
		(0.152)				(0.095)		
$Knowledge_{it}^{Up,IO}$			0.258				0.198	
			(0.165)				(0.203)	
Specification	OLS	OLS	OLS	IV 2nd Stage	OLS	OLS	OLS	IV 2nd Stage
IV 1st Stage F-statistics				427				427
R^2	0.916	0.917	0.917	0.169	0.900	0.900	0.900	0.092
No. of Sectors	95	95	95	95	95	95	95	95
No. of Obs	1900	1900	1900	1140	1900	1900	1900	1140
Fixed Effects	Sector, Year				Sector, Year			

Notes. This table tests the relation between innovation in a focal sector and past innovation in sectors connected through the innovation network, using the U.S. data from BLS sectors. We restrict the sample to sectors that have at least 100 patents over the full sample period. Standard errors in parentheses are clustered at the sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 2, column (1) presents the results of regression (27). Sectoral R&D expenditure significantly predicts the number of new patents filed in a given year, with an elasticity of 0.275. The knowledge stock upstream of each sector (Knowledge$_{it}^{Up}$) also significantly predicts patent output, with an elasticity of 0.586. Column (5) shows that both variables also predict patent quality: sectors with greater R&D and greater upstream knowledge stock tend to produce patents with more future citation counts. In the Online Appendix Table A.11, we also demonstrate these variables predict the commercial value of innovation measured using the stock market reaction upon patent approval (Kogan et al., 2017).

These regressions provide supportive evidence that past knowledge in sectors upstream of i benefits subsequent patent production in the focal sector i. An alternative story relates to common shocks: a group of sectors connected to each other via citation linkages may face similar demand, supply and investment opportunities, leading to co-movements of innovation activities. Serial correlations in such common shocks would lead to a positive coefficient β_1 in regression (27) even without cross-sector knowledge spillovers. This is a version of the "reflection problem" à la Manski (1993) and Bloom et al. (2013).

We implement three additional analyses to address the "common shock" concern. First, we exploit the directional nature of knowledge spillovers. We construct the knowledge stock aggregator for sectors *downstream* of i in the innovation network:

$$\text{Knowledge}_{it}^{Down} \equiv \sum_{k \neq i} \sum_{\tau=1}^{10} \omega_{ki,t-\tau} \ln n_{kt-\tau}.$$

$\text{Knowledge}_{it}^{Down}$ aggregates the patent output in all sectors $k \neq i$, weighted by the extent to which patents in sector k cite those in sector i, and is therefore a measure of the knowledge stock *downstream* of sector i. Because knowledge flow is directional, there should be an asymmetry: while $\text{Knowledge}_{it}^{Up}$ should positively predict subsequent patent output in sector i, $\text{Knowledge}_{it}^{Down}$ should not. Yet any common shocks hitting this network should generate symmetric correlations in innovation output for focal sector i and both its upstream and downstream sectors.

Columns (2) and (6) of Table 2 add $\text{Knowledge}_{it}^{Down}$ to our baseline regressions. We make two observations. First, adding $\text{Knowledge}_{it}^{Down}$ as a control does not meaningfully affect the economic or statistical significance of our two baseline variables. This suggests our baseline regressions are not simply picking up correlated shocks to local technology clusters. Second, the coefficient on $\text{Knowledge}_{it}^{Down}$ is precisely zero, confirming our key model mechanism and that knowledge flow along the innovation network is directional—it goes only from upstream to downstream, and not the other way around.

Another related concern is that common shocks operate not through technological linkages but through input-output (IO) linkages. To address this, we construct the aggregator $\text{Knowledge}_{it}^{Up,IO}$ similarly to $\text{Knowledge}_{it}^{Up}$, but using patents from other sectors weighted not by the innovation network, as in (26), but instead by sector i's cost share on inputs from sector j. Columns (3) and (7) of Table 2 show the regression results when including $\text{Knowledge}_{it}^{Up,IO}$. Knowledge from innovation-upstream sectors remains an economically and statistically significant predictor of subsequent innovation in the focal sector. By contrast, $\text{Knowledge}_{it}^{Up,IO}$ has a smaller impact on sector i's patent quantity, and insignificant effect on innovation quality in these specifications. These results, along with the fact that the innovation network only weakly correlates with the IO network (see Table 1), imply that the innovation network provides valuable incremental information that is particularly powerful for understanding knowledge spillovers across sectors.

Next, we adopt another approach to address the "common shock" concern using tax-induced changes to the effective cost of R&D to create exogenous variations in innovation activities, following Bloom, Schankerman, and Van Reenen (2013). We briefly describe the approach here and provide more details in the Online Appendix E.3. The approach leverages the fact that the user-cost of R&D capital (i.e., the cost of conducting R&D) varies with state-level R&D tax credit, depreciation allowance, and corporate tax rate. Cross-sector heterogeneity in the geographic distribution of R&D activities in turn translates into R&D cost differences across sectors and over time. For the U.S., we use Wilson (2009)'s estimates of state-specific R&D cost shifters, combined with our estimates of the cross-state distribution of each sector's R&D, to calculate a sector's R&D

costs. For the global setting below, we follow Thomson (2017) to calculate the user cost of R&D capital at the country-sector-year level.

We first create fitted values of sectoral innovation output ($\ln n_{it}$) using R&D cost shifters; we then use these fitted values in equation (26) to construct a predicted value of $\widehat{\text{Knowledge}}_{it}^{Up}$, which is in turn used as an instrumental variable (IV) for $\text{Knowledge}_{it}^{Up}$ for a two-stage least-squares (2SLS) analysis. Columns (4) and (8) in Table 2 remain qualitatively and quantitatively robust to using this IV strategy. Details of this analysis are provided in Online Appendix E.3.

4.2.2 Global Evidence

We now test international knowledge spillovers in our global sample. First, we construct an analogous measure of upstream knowledge stock: for each focal country m, sector i in year t, we enumerate over all countries c and sectors j in our sample, aggregating the (log-)patent output in cj over the past 10 years, weighted by the share of mi's citations that are to cj in the corresponding year:

$$\text{Knowledge}_{mit}^{Up} \equiv \sum_{cj \neq mi} \sum_{\tau=1}^{10} \frac{Cites_{mi \to cj, t-\tau}}{\sum_{c'=1}^{N} \sum_{k=1}^{K} Cites_{mi \to c'k,, t-\tau}} \ln n_{cj,t-\tau}. \tag{28}$$

Next, we adapt our closed economy test of knowledge spillovers to perform on the global innovation network. In this case, the unit of observation is at the country-industry-year level, and we include a saturated set (country-industry, country-year, industry-year) of fixed effects:

$$\ln n_{mit} = \beta_1 \times \text{Knowledge}_{mit}^{Up} + \beta_2 \times \ln R\&D_{mi,t-1} + \xi_{mi} + \xi_{mt} + \xi_{it} + \varepsilon_{ict}. \tag{29}$$

Table 3 shows the results: knowledge stock upstream of each country-industry significantly predicts subsequent patent counts (column 1) and citation-adjusted patent counts (column 5) even in the global setting. The coefficients are lower than estimates based only on the U.S., suggesting that knowledge spillovers are stronger across sectors within the U.S. than they are across countries, potentially due to barriers to cross-border knowledge diffusion such as cultural and language differences, and inappropriateness of foreign technology.

To rule out common shocks to technological and input-output clusters, we again make use of knowledge aggregated from downstream, from the input-output network, and the tax-induced IV, and find evidence in support of knowledge spillovers through the innovation network.

5 Application: R&D Resource Allocation in the Data

This section hosts our main empirical analysis, which uses our model to study the allocation of R&D resources in the data. We present the computed unilaterally optimal allocation of R&D resources across sectors for each country in our sample. We show that on average, sectors that

Table 3. Global Evidence of Knowledge Spillovers Through the Innovation Network

$Y=$	ln(Patents)				ln(Cites)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Knowledge^{Up}_{mit}$	0.162***	0.188***	0.159***	0.226**	0.352***	0.393***	0.350***	0.453***
	(0.055)	(0.056)	(0.055)	(0.113)	(0.077)	(0.080)	(0.078)	(0.143)
$ln(R\&D\ Stock)_{mi,t-1}$	0.043***	0.043***	0.043***	0.079***	0.084***	0.084***	0.083***	0.083***
	(0.013)	(0.013)	(0.013)	(0.020)	(0.018)	(0.018)	(0.018)	(0.030)
$Knowledge^{Down}_{mit}$		-0.059				-0.094		
		(0.039)				(0.062)		
$Knowledge^{Up,IO}_{mit}$			0.070				-0.054	
			(0.065)				(0.068)	
Specification	OLS	OLS	OLS	IV 2nd Stage	OLS	OLS	OLS	IV 2nd Stage
IV 1st Stage F-statistics				148				148
R^2	0.968	0.968	0.968	0.035	0.943	0.943	0.943	0.028
No. of Country x Sectors	570	570	556	282	570	570	556	282
No. of Obs	11014	11014	10774	4587	11014	11014	10774	4587
Fixed Effects	Country x Sector, Country x Year, Sector x Year				Country x Sector, Country x Year, Sector x Year			

Notes. This table tests the relation between innovation in a focal sector and past innovation in connected sectors through the innovation network, in an international setting. We restrict the sample to country-sectors with at least 10 patents over the full sample period. Standard errors in parentheses are clustered at the country-sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

should have more R&D resources do receive more resources, especially for the five economies with the most patents during our sample period. Nevertheless, the residual misalignment between the optimal and actual allocations remains large and is highly heterogeneous across countries. We compute the welfare gains from adopting the optimal R&D allocation for each country.

5.1 Optimal R&D Allocation

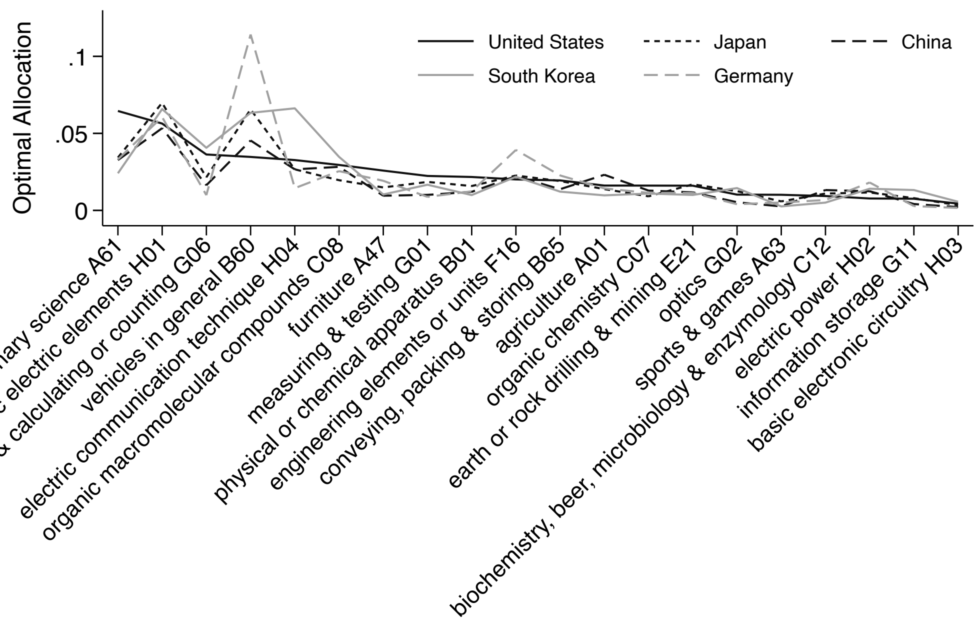
For each country and year, we calculate the unilaterally optimal cross-sector allocation of R&D resources $\boldsymbol{\gamma}$, using Proposition 6:

$$\boldsymbol{\gamma}' = \xi^{-1} \frac{\rho}{\rho + \lambda} \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega} \circ \boldsymbol{X}}{1 + \rho/\lambda} \right)^{-1}, \tag{30}$$

where the proportionality constant ξ ensures that elements in the optimal allocation vector $\boldsymbol{\gamma}$ sum to one and is a measure of R&D self-sufficiency. We measure $\boldsymbol{\beta}$ using each country's sectoral value-added relative to GDP in that year (thereby accounting for input-output linkages; see Section 2.7.1). Recall that $\boldsymbol{X} \equiv [x_{ij}]$ is the matrix encoding the share of domestic contribution to cross-sector knowledge spillovers; we measure x_{ij} as the share of citations from i to sector j that are toward domestic patents in j. As Figure 3 shows, entries of $\boldsymbol{X}$ average to above 70% across sectors for the U.S. but are significantly lower for all other countries except Japan in recent years. For $\boldsymbol{\Omega}$, we use the innovation network built using all global patents filed within ten years up to the year prior to the analysis. That is, for analysis of 2010, we use $\boldsymbol{\Omega}$ constructed using all patents filed between 2000 and 2009.

To implement the formula in (30), we need to specify the discount rate ρ and the elasticity λ of knowledge growth with respect to the new innovation flows. As a baseline, we set discount rate $\rho = 5\%$. Because knowledge stock does not have a natural scale, we rely on the empirical pattern of knowledge spillover dynamics to calibrate λ. Specifically, equation (25) implies that the spillover effect of past upstream knowledge on current innovation decays over time at rate λ. We estimate the spillover effects of past upstream patents nonparametrically at different time lags (Appendix Figure A.9). We find that the knowledge spillovers have a half-life of about four years, corresponding to $\lambda = 0.17$, which we adopt as the baseline calibration. Qualitatively, our analysis is not sensitive to the value of ρ/λ: as we show in the Online Appendix Table A.20, the optimal allocation $\boldsymbol{\gamma}$ is highly correlated across specifications with alternative values of ρ/λ.[5] The Online Appendix also reports additional results and sensitivity checks such as using data from other sample periods, using alternative specifications for $\boldsymbol{\Omega}$, and calibrate sector-specific λ_i. We describe these and other extensions in Section 5.4.

Figure 4. Optimal R&D Allocations in Different Countries



Notes. This figure shows the optimal R&D allocation across 20 3-digit IPC classes with the most patents for the five economies that produced the most patents in 2010–2014. Optimal allocations are calculated using our baseline calibration $\rho = 5\%$,$\lambda = 0.17$. Sectors are sorted by the optimal allocation for the U.S.,γ_{US}.

Figure 4 plots the optimal R&D allocation $\boldsymbol{\gamma}$ for the five economies that produced the most

[5]One reason why our conclusions are robust to alternative values of ρ/λ is that, as discussed previously, an increase in ρ/λ has the same implication for the optimal R&D allocation as an increase in a country's reliance on foreign knowledge. The substantial cross-country variation in foreign reliance (see Figure 3) dwarfs reasonable variation in the calibration of ρ/λ. Hence, the qualitative cross-country differences in unilaterally optimal R&D allocations are not sensitive to our calibration of ρ/λ.

patents in 2010–2014. For visual clarity, we only show the top 20 3-digit IPC classes ranked by total patent counts; these 20 classes account for 75% of all patents. The level of optimal R&D resources is shown on the y-axis, and the x-axis represents IPCs (sorted by γ_{US}).

For the U.S. (solid black line), the optimal allocation favors sectors with the highest innovation centrality, as listed in Figure 2, such as medical science (A61), basic electric elements (H01, e.g., semiconductors), and computing devices (G06). The top 10 IPCs (out of 131) should receive about a third of total U.S. R&D resources. The correlation between optimal U.S. R&D allocation and the innovation centrality $\boldsymbol{a}$ is 0.75. The correlation is high because the U.S. has a self-contained innovation network with relatively few citations toward foreign patents; hence, its planner should internalize more knowledge spillovers. The correlation is not perfect since the planner also considers IPC's importance for domestic production, encoded in the value-added share vector $\boldsymbol{\beta}$, which raises the optimal allocation of high-β sectors such as vehicles (B60).

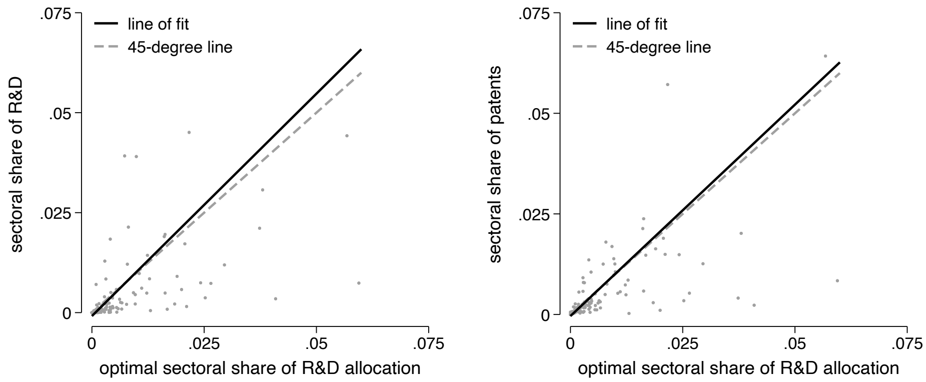
Figure 4 also reveals cross-country variations in the unilaterally optimal R&D allocations. Relative to the U.S., Germany and Japan should allocate more toward vehicles (B60); South Korea should allocate more toward electric communication technique (H04); all four non-U.S. economies should allocate less toward medical science (A61).

5.2 Innovation Allocation in the Data

We first present our model's ability to fit R&D resource allocation in the U.S. The left panel of Figure 5 shows the scatter plot of sectoral R&D expenditure (as a share of total R&D) against the optimal R&D expenditure share γ_{US} for the sample period 2010–2014. The linear fit (solid line) is close to the 45-degree line (dashed) with a slope of 1.11 (t-statistic 7.64), indicating that on average, sectors that should optimally receive more R&D resources do indeed receive more R&D resources. In the right panel of Figure 5, we change the y-axis to sectoral patent output as a share of total patent output; again, sectoral patent output aligns very well with γ_{US}, with a slope of 1.05 (t-statistic 8.20). To be clear, the strong alignment between real-world and optimal R&D allocations does not imply the U.S. allocates R&D optimally: there is substantial residual variation in R&D allocations as they disperse around the 45-degree line. The vertical distance between each observation and the 45-degree line measures the amount of R&D resources that need to be reassigned to achieve the optimal allocation. We quantitatively assess the welfare gains from adopting the optimal R&D allocation in Section 5.3 below.[6]

[6] A potential concern is that Figure 5 picks up a mechanical relationship: it may be that sectors with more resources produce more patents and citations, thereby appearing to be more central in the innovation network $\boldsymbol{\Omega}$—in other words, allocated resources reversely affect sectoral centrality. To argue against this possibility, we reproduce our empirical exercises using the innovation network constructed using citations from Japanese patents to Japanese patents. Because Japan's innovation network is self-contained and has few citations toward foreign patents (Figure 3), the network is by construction independent of U.S. R&D. All of our findings continue to hold, suggesting that in-

Figure 5. U.S. Actual R&D Allocation vs. Optimal Allocationγ_{US}



Notes. This figure shows scatter plots of real-world sectoral R&D expenditure shares (left panel) and patent output shares (right panel) against optimal R&D allocation shares, γ_{US}, for the U.S. in 2010-2014. The solid line is the linear fit; the dashed line is the 45-degree line. For visual clarity, we exclude 3 outlier sectors (out of 131) that account for >7.5% of R&D shares or patent output from the scatter plots, but all sectors are used for the linear fit.

There is substantial cross-country heterogeneity in R&D resource allocations. Figure 6 shows scatter plots of sectoral R&D expenditure shares against the unilaterally optimal R&D allocations for ten countries that filed the most patents in 2010–2014. Sectoral R&D expenditure correlates strongly with the optimal R&D allocations for the five countries shown in the top row (U.S., Japan, China, South Korea, and Germany), and the relationship is weaker for the five economies at the bottom (Russia, France, U.K., Canada, and Netherlands). As we have noted, a line-of-fit with a slope of 1 does not imply resources are allocated optimally; nevertheless, Figure 5 suggests that on average, more resources need to be reallocated to achieve optimality in the five economies in the bottom row. These results are robust to using patent output shares and R&D shares reported in OECD ANBERD database as measures of cross-sector R&D resource allocation (Figures A.13 and A.15 in the Online Appendix), suggesting that our results are not driven by the coverage and quality of our R&D expenditure variable. Figure A.12 in the Online Appendix shows that very similar patterns hold in the years 2000 and 2005.

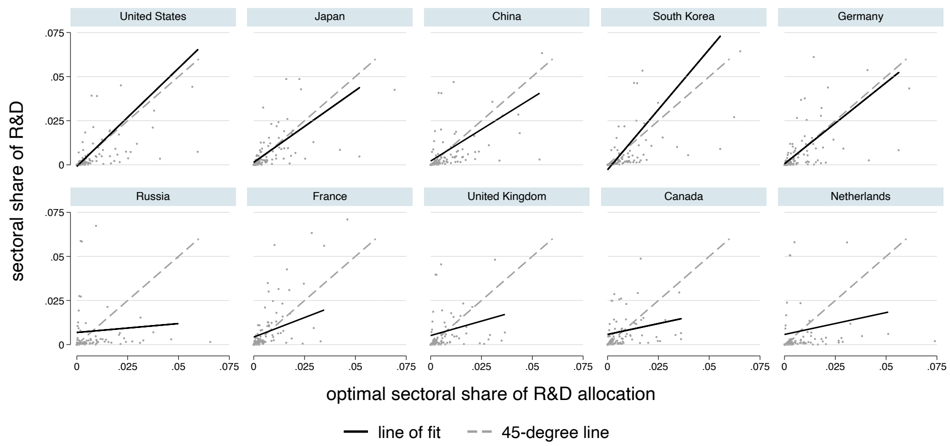
5.3 Welfare Gains from Improving R&D Allocation

We now quantify the potential welfare gains from improving R&D allocation, using the welfare formula in Proposition 7:

$$\ln \mathcal{L}(\boldsymbol{b}, \xi) = \psi \frac{\lambda}{\rho} \underbrace{\xi}_{\text{self sufficiency}} \times \underbrace{\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})}_{\text{R\&D misallocation}}, \tag{31}$$

novation centrality $\boldsymbol{a}$—which correlates strongly with the U.S. optimal R&D allocation $\boldsymbol{\gamma}^{US}$—indeed picks up sectoral importance in the innovation network rather than representing historical R&D expenditures.

Figure 6. Actual R&D Allocation vs. Optimal Allocation Across Countries



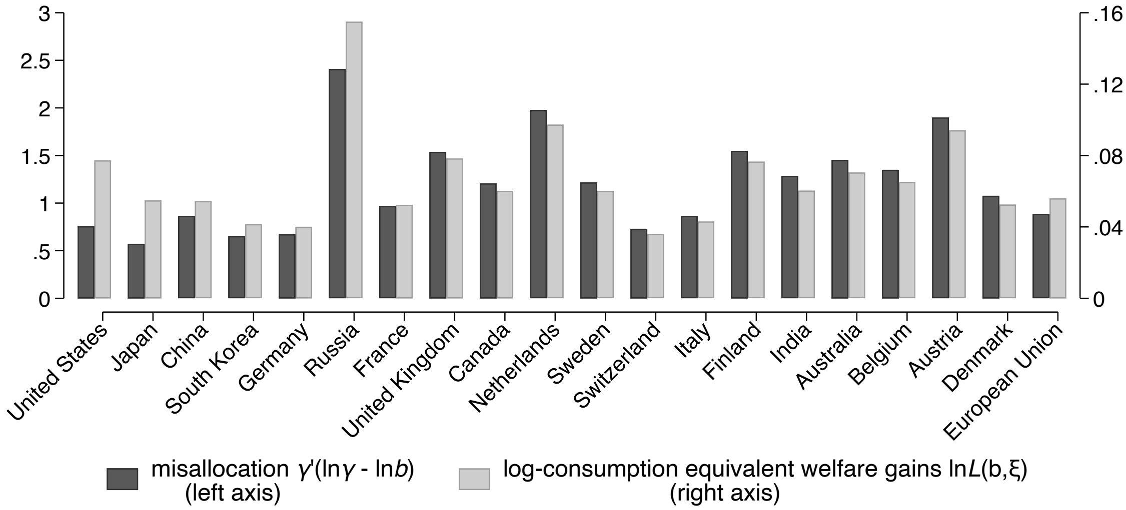
Notes. This figure shows scatter plots of sectoral R&D expenditure shares against the optimal sectoral share of R&D allocation for the top ten innovative countries (in terms of patent output) using data from 2010–2014. The solid line is the linear fit; the dashed line is the 45-degree line. For visual clarity, we exclude outlier sectors that account for >7.5% of national R&D shares from the scatter plots, but all sectors are used for the linear fit.

where $\ln \mathcal{L}(\boldsymbol{b}, \xi)$ is the consumption-equivalent welfare gains in logs, $\boldsymbol{b}$ is the empirical R&D allocation vector, $\boldsymbol{\gamma}$ is the optimal R&D allocation vector, and $\xi \equiv \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^{-1}\mathbf{1}$ is the scalar measure of R&D self-sufficiency, which increases in domestic citation shares x_{ij}. ψ captures the elasticity of productivity to knowledge stock and thus proportionally controls the welfare impact of R&D allocation. It is an additional parameter to be calibrated. We specify $\psi = 0.06$, which implies a semi-elasticity $(\,\mathrm{d}g^y/\,\mathrm{d}\ln\bar{s} = \psi\lambda =)$ 0.01 of BGP consumption growth rate to the total stock of R&D resource, consistent with standard calibrations in the growth literature (Akcigit and Kerr, 2018, Akcigit et al., 2021).

Equation (31) implies a natural decomposition when comparing across countries the potential welfare gains from reallocating R&D: misallocation, measured as relative entropy from the actual allocation $\boldsymbol{b}$ to the optimal $\boldsymbol{\gamma}$ (higher $\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b})$ means less efficient) and self-sufficiency (ξ). Holding an economy's R&D self-sufficiency constant, worse misallocation implies larger potential gains; and, as an economy benefits more from foreign knowledge spillovers (lower ξ), domestic R&D misallocation becomes less consequential for consumer welfare.

Figure 7 shows the degree of misallocation and the size of the potential welfare gains from R&D reallocation for the 19 economies that filed the most patents during 2010–2014. Because of the EU's high degree of economic integration, we also aggregate R&D from all EU countries and calculate the allocative efficiency of the EU as a single, integrated economy, which is listed as the 20th economy in Figure 7. The dark bars represent the misallocation term (left Y-axis). Among high-income countries, Japan has the most efficient R&D allocation, followed by other top-patenting

Figure 7. R&D Allocative Efficiency and Potential Welfare Gains Across Countries



Notes. This table shows the level of R&D misallocation (dark bars) and the potential welfare gains (light grey bars) from adoption optimal R&D allocation across 19 innovative countries with the highest patent outputs in our sample, and the integrated economy of the European Union, using 2010-2014 data. The calculation focuses on R&D in top 50 IPC classes by total patents.

countries like the U.S., Germany, and South Korea. Among high-income economies in Europe, Switzerland also has a higher allocative efficiency compared to its peers. When evaluated as a single integrated economy, the EU's allocative efficiency is comparable to Italy's and is marginally better than France's.

The light grey bars in Figure 7 represent the potential welfare gains of reallocating R&D optimally (right Y-axis). By our welfare accounting formula (31), the welfare gain is proportional to the misallocation term times the R&D self-sufficiency measure ξ. For the two economies with self-contained innovation networks, namely the U.S. and Japan, ξ is closer to one, so the overall welfare gains (grey bars) are closer in magnitude to the misallocation terms (dark bars). By contrast, the welfare gains are comparatively lower than the corresponding misallocation terms in all other economies, as their domestic R&D misallocation is less consequential for economic growth because of their dependence on foreign knowledge spillovers.

Table 4 shows the size of the potential welfare gains from adopting the optimal allocation, evaluated using R&D allocation measured in the years 2000, 2005, and 2010. For the year 2010, adopting the optimal allocation in Japan, which has the most efficient R&D allocation in our sample, could lead to welfare improvements equivalent to raising consumption along the entire path by 5.64%. The potential welfare gains for the U.S. are 8.04% in consumption-equivalent terms,[7] which is above the average in our sample. Russia has the highest potential gains, equivalent to

[7] In Section B.6 of the Online Appendix, we derive the optimal R&D allocation and the welfare cost formula when the domestic planner takes into account how domestic R&D affects foreign variables. We find very similar welfare gains under that specification.

16.76% consumption gains after adopting the optinal allocation. Moving to the country-specific optimal R&D allocation in 2010 would generate consumption-equivalent welfare gains of 5.60% in China, 4.24% in South Korea, and 4.09% in Germany. For most economies, the size of the potential welfare gains has been stable since the 2000s.

It is important to note that a more allocatively efficient country is not necessarily more innovative in absolute terms. Instead, the extent of misallocation reflects the distance from actual R&D allocation ($\boldsymbol{b}$) to each country's own efficient benchmark ($\boldsymbol{\gamma}$), and our welfare calculations reflect how much each country could gain when moving to that benchmark, holding all other economic conditions fixed.

Table 4. Percentage (%) Consumption Gains By Moving to Each Country's Optimal R&D Allocation

	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
2000	9.98	4.24	5.78	5.25	4.79	13.70	5.17	7.55	7.22	6.70
2005	8.85	5.04	5.26	3.92	4.11	11.18	5.38	8.17	7.29	5.45
2010	8.04	5.64	5.60	4.24	4.09	16.76	5.38	8.15	6.21	10.22

	Sweden	Switzerland	Italy	Finland	India	Australia	Belgium	Austria	Denmark	European Union
2000	6.65	5.18	5.04	5.39	10.91	5.72	5.72	6.52	5.93	5.91
2005	5.53	4.10	4.57	5.63	8.33	4.19	5.62	8.50	5.30	5.04
2010	6.20	3.67	4.40	7.95	6.21	7.30	6.73	9.87	5.39	5.76

Notes. This table shows the consumption-equivalent welfare gains when each economy moves its optimal R&D allocation, calculated using the formula (31) under our baseline calibration $\{\rho, \lambda, \psi\} = \{0.05, 0.17, 0.06\}$. Section 5.4 discusses robustness under different parameter values.

5.4 Additional Results, Robustness Checks, and Sensitivity Analysis

In Online Appendix E.4, we present additional results and robustness checks on our quantitative analysis of R&D reallocation. Note that our R&D allocation accounting exercise rely on information on the network $\boldsymbol{\Omega}$, the value of the effective discount rate ρ/λ, and the elasticity of productivity to knowledge ψ. In the baseline specification, we compute entry ω_{ij} of $\boldsymbol{\Omega}$ as the share of all citations from i that are towards j ($\frac{Cites_{i\to j}}{\sum_k Cites_{i\to k}}$), and we set $\{\rho, \lambda, \psi\} = \{0.05, 0.17, 0.06\}$. For expositional simplicity, the baseline results focus on top innovative sectors and in the most recent time period. We now tackle the robustness along all these dimensions in the Online Appendix.

Table A.20 shows that the optimal R&D allocation $\boldsymbol{\gamma}$ is highly stable (by both Pearson's correlation and Spearman's rank correlation) across the following alternative specifications of the innovation network $\boldsymbol{\Omega}$ and parameterizations of ρ and λ.[8]

[8] We do not report robustness results under alternative values of parameter ψ, as the parameter does not affect the optimal allocation $\boldsymbol{\gamma}$. The welfare impact of R&D allocation is directly proportional ψ, so our baseline results in Table 4 can be directly rescaled for different values of ψ.

As discussed in Section 4.1, cross-sector knowledge spillovers are inherently difficult to capture. We consider several alternative specifications of $\mathbf{\Omega}$. In rows A1 and A2, we weigh each citation linkage in $\mathbf{\Omega}$ construction (24) by the quality of either the citing (row A1) or the cited patent (row A2) measured using the total forward citations received by these patents. This way, we over-weigh the citation linkage when the technologies involved are impactful. The innovation networks in these extensions therefore weigh more heavily the spillovers among major patents. In row A3, we construct $\omega_{ij} \propto Cites_{i\to j}$ to scale directly with the total citations totally across or ij-pairs (rather than normalized by the citations from i), and we choose the proportionality constant so that the spectral radius of $\mathbf{\Omega}$ is equal to one, ensuring endogenous growth as in our baseline model (see Section B.9 for the theoretical discussion of this specification).

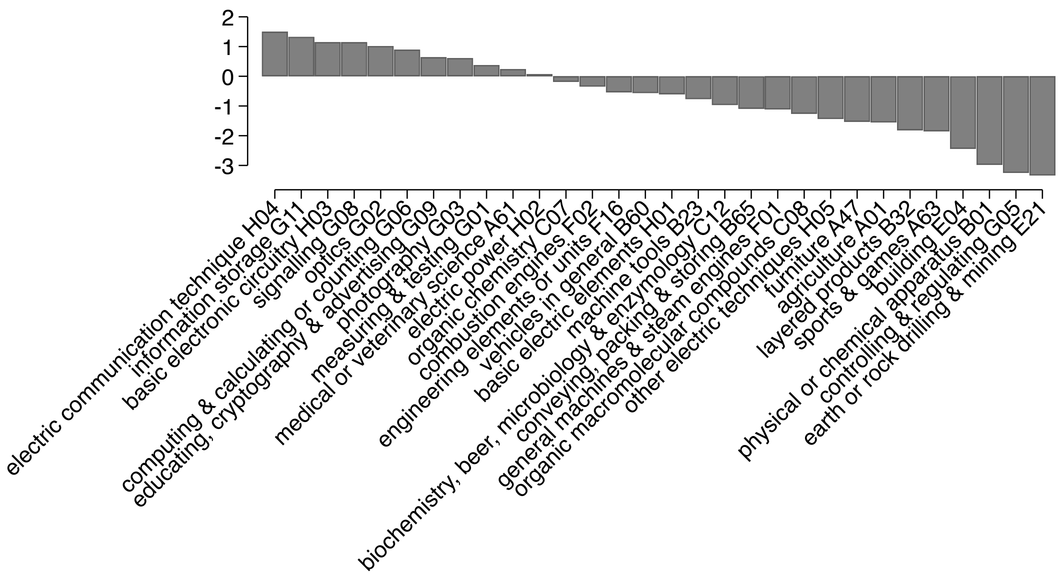
In rows B1 to B7, we consider a range of alternative values for ρ/λ and show that the optimal R&D allocation correlates highly with our baseline specification. We also consider a specification with sector-specific λ_i (row C1). The optimal R&D allocation in this environment is derived in Online Appendix B.8; the heterogeneity in λ_i is measured using variations in each sector's median ROA (return on assets) in our firm-level datasets, with the mapping motivated by the decentralized economy constructed in Online Appendix B.4.

Finally, one related concern is that citations are only noisy proxies of knowledge spillovers, and thus the innovation network can only be noisy measured. While we cannot purge the measurement error, we conduct the robustness check in reverse and show our quantitative analysis is robust to introducing additional, simulated random errors to $\mathbf{\Omega}$, shown in rows D1 to D10.

5.5 How Does R&D Allocation Compare with the Optimal in the U.S.?

We here provide some descriptive evidence for how the actual R&D allocation compares with the optimal in the U.S. Figure 8 plots the log-ratio between the actual R&D expenditure share in the U.S. and the optimal allocation for the 30 largest 3-digit IPC classes by patent output. Altogether, these 30 IPC classes (out of 131) account for 84% of patents and 90% of R&D expenditures in the U.S. Though providing a full set of policy recommendations on R&D allocation is beyond the scope of this paper, this figure conveys several noteworthy messages. Electric communication technique (H04; e.g., telephonic communication, wireless communication), which ranks 4th in centrality (Figure 2) and 9th in the optimal allocation γ (Figure 4), is over-invested. Meanwhile, within the same broad IPC class H (electricity), the more central and fundamental class Basic Electric Elements (H01; e.g., semiconductor devices) is underinvested. In terms of its economic magnitude, -0.60, our log-ratio scale suggests that the real allocation is 55% ($\exp(-0.60)$) of the optimal level; in other words, R&D in semiconductor devices is underfunded by about 45%. This supports the recent U.S. initiatives (such as the CHIPS For America Act) to invest in the semiconductor industry. Another observation is that the underinvested group (right end of the graph) over-represents IPC classes related to technologies often termed as "green innovation" (see Cohen, Gurun, and Nguyen 2020) that can help reduce pollution and the negative consequences of resource exploitation. For instance, one of the most underinvested IPC class in the figure, B01, covers subclasses on waste management, alternative energy production, and environmental management. Figure A.17 of the Online Appendix further demonstrates the R&D allocative efficiency in the US across 3-digit IPC classes within each broad 1-digit IPC technology class.

Figure 8. U.S. R&D Misallocation in the Top 30 Innovative IPC Classes



Notes. This figure plots the level of misallocation of the top 30 innovative IPC classes, ranked using total patent output. The level of misallocation is calculated as $\ln \boldsymbol{b} - \ln \boldsymbol{\gamma}$. Positive bars (left end) imply over-investment, and negative bars imply underinvestment.

5.6 Innovation Hubs

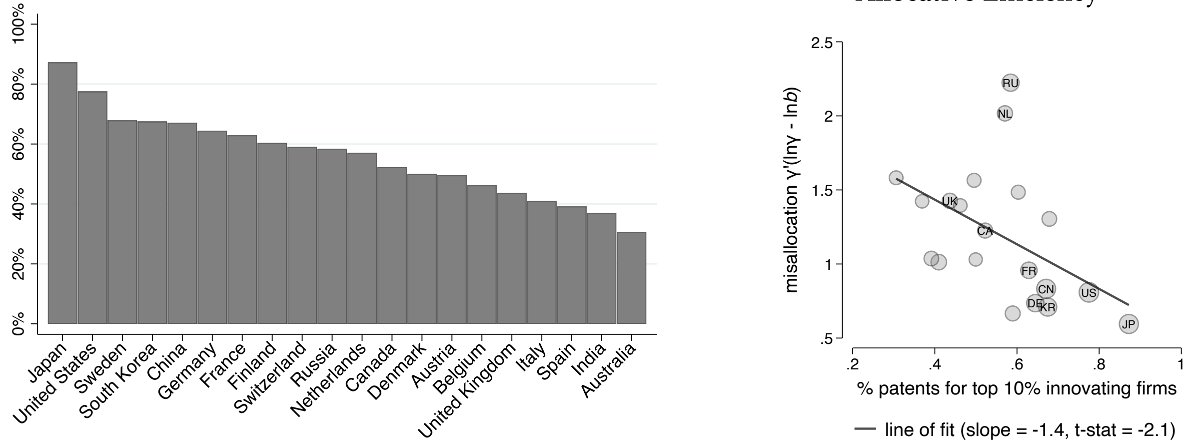
What explains cross-country differences in R&D allocative efficiency? We do not have definitive answers, but we can present a conjecture with some empirical support: firms whose R&D activities span multiple sectors and technology classes allocate their resources in ways that may resemble the social planner's. Because these firms' R&D activities build on their own prior innovations, they may partially internalize knowledge spillovers through the innovation network. Notable examples include top innovating firms such as IBM, Samsung, Sony, and Siemens, which are termed "innovation hubs."

Our hypothesis is supported empirically by a strong negative relationship between the presence of such firms and the degree of R&D misallocation in each country. Figure 9, Panel (a) shows the share of patents in 2010–2014 that are filed by the top 10% of innovating firms in each country.

Figure 9. Innovation Hubs and R&D Allocative Efficiency

(a) Share of Patents from Top Innovative Firms

(b) Innovation Hubs and R&D Allocative Efficiency



Notes. Panel (a) of this figure shows the share of patents filed by the top 10% of innovative firms in each country between 2010–2014 (innovative firms are ranked using patent output). Panel (b) plots the misallocation measure against the measure of concentration in Panel (a).

The figure shows that R&D activities are more concentrated in Japan, the U.S., and Sweden, as the top 10% of innovating firms in these economies account for close to 90%, 80%, and 70% of patents, respectively. By contrast, R&D activities are least concentrated in Spain, India, and Australia.

Panel (b) of Figure 9 plots the misallocation measure ($\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln \boldsymbol{b})$) against the share of patents accounted for by the top 10% of innovating firms. We find a strongly negative relationship (slope -1.4, t-statistic -2.1). This evidence suggests that the market failure in R&D resource allocation could be partially mitigated if innovation hub firms thrive.

6 Conclusion

We study optimal cross-sector allocation of R&D resources in an endogenous growth model featuring an innovation network. We provide closed-form solutions for the optimal path of R&D resource allocation, and we show a planner valuing long-term growth (i.e., with low discount rates) should allocate more R&D toward key sectors that are central in the innovation network, but the incentive is muted in open economies that benefit more from foreign knowledge spillovers. We show the relative entropy of actual R&D allocation from the optimal allocation maps into a sufficient statistic for the potential welfare gains from reallocating R&D optimally.

To empirically evaluate R&D allocative efficiency across countries and over time, we build a global innovation network based on over 30 million global patents and compile comprehensive data on sectoral production, final use, and, importantly, R&D resource allocation for major innovative economies. We find that our model-implied optimal R&D resource allocation explains real allocations in the data, particularly for countries generally perceived as innovative, such as the U.S., Japan, Germany, and more recently China and South Korea. However, there remains significant room to improve. Improving R&D allocations could generate substantial welfare improvements across the globe. For the U.S., reallocating R&D resources to Japan's efficiency level would increase consumption-equivalent welfare by 19.6% in 2010. We believe our framework can be adopted to explore future questions about R&D allocation.

References

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2018): "Innovation, reallocation and growth," *American Economic Review*, 108, 3450–3491.

Acemoglu, D., U. Akcigit, and W. Kerr (2015): "Networks and the Macroeconomy: An Empirical Exploration," 273–335.

Acemoglu, D., U. Akcigit, and W. R. Kerr (2016): "Innovation network," *Proceedings of the National Academy of Sciences*, 113, 11483–11488.

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005): "Competition and innovation: An inverted-U relationship," *Quarterly Journal of Economics*, 120, 701–728.

Akcigit, U., S. Ates, and G. Impullitti (2022): "Innovation and Trade Policy in a Globalized World," *Working paper.*

Akcigit, U., D. Hanley, and N. Serrano-Velarde (2021): "Back to basics: Basic research spillovers, innovation policy, and growth," *Review of Economic Studies*, 88, 1–43.

Akcigit, U. and W. R. Kerr (2018): "Growth through heterogeneous innovations," *Journal of Political Economy*, 126, 1374–1443.

Atkeson, A. and A. Burstein (2019): "Aggregate implications of innovation policy," *Journal of Political Economy*, 127, 2625–2683.

Baqaee, D. R. (2018): "Cascading Failures in Production Networks," *Econometrica*, 86.

Baqaee, D. R. and E. Farhi (2019): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," *Econometrica*, 87.

——— (2020): "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020): "Are ideas getting harder to find?" *American Economic Review*, 110, 1104–44.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying technology spillovers and product market rivalry," *Econometrica*, 81, 1347–1393.

Bloom, N., J. Van Reenen, and H. Williams (2019): "A toolkit of policies to promote innovation," *Journal of Economic Perspectives*, 33, 163–84.

Buera, P. and E. Oberfield (2020): "The Global Diffusion of Ideas," *Econometrica.*

Caballero, R. J. and A. B. Jaffe (1993): "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," *NBER Macroeconomics Annual*, 8, 15–74.

Cai, J. and N. Li (2019): "Growth Through Intersectoral Knowledge Linkages," *Review of Economic Studies.*

Cai, J., N. Li, and A. M. Santacreu (2022): "Knowledge Diffusion, Trade and Innovation across Countries and Sectors," *AEJ: Macroeconomics*, 14, 104–45.

Cai, J. and C. Tian (2021): "Innovation and Endogenous Knowledge Network Dynamics," *Working paper.*

Carvalho, V. (2010): "Aggregate fluctuations and the network structure of intersectoral trade," *Working paper.*

Chaney, T. (2018): "The Gravity Equation in International Trade: An Explanation," *Journal of Political Economy.*

Coe, D. T. and E. Helpman (1995): "International R&D spillovers," *European Economic Review.*

Coe, D. T., E. Helpman, and A. W. Hoffmaister (2009): "International R&D Spillovers and Institutions," *European Economic Review.*

Cohen, L., U. G. Gurun, and Q. H. Nguyen (2020): "The ESG-innovation disconnect: Evidence from green patenting," .

David, J. and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review.*

Eaton, J. and S. Kortum (1999): "International Technology Diffusion: Theory and Measurement," *International Economic Review.*

——— (2006): "Innovation, Diffusion, and Trade," *NBER Working paper.*

Gabaix, X. (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica.*

Galeotti, A., B. Golub, and S. Goyal (2020): "Targeting Interventions in Networks," *Econometrica.*

Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow (2019): "How Destructive Is Innovation?" *Econometrica*, 87, 1507–1541.

Grassi, B. (2017): "IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important," *Working Paper.*

Griffith, R., R. Harrison, and J. Van Reenen (2006): "How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing," *American Economic Review*, 96, 1859–1875.

Guillard, C., R. Martin, P. Mohnen, C. Thomas, and V. Dennis (2021): "Efficient industrial policy for innovation: standing on the shoulders of hidden giants," *Working paper.*

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005): "Market value and patent citations," *RAND Journal of Economics*, 16–38.

Hopenhayn, H. and F. Squintani (2021): "On the Direction of Innovation," *Journal of Political Economy.*

Hsieh, C.-T., E. Hurst, C. Jones, and P. Klenow (2019): "The Allocation of Talent and U.S. Economic Growth," *Econometrica.*

Hsieh, C.-T. and P. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics.*

Huang, J. and Y. Zenou (2020): "Key Sectors in Endogeneous Growth," *Working paper.*

Hulten, C. R. (1978): "Growth Accounting with Intermediate Inputs," *The Review of Economic Studies*, 45, 511–518.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics*, 108, 577–598.

Jones, B. F. and L. H. Summers (forthcoming): "A Calculation of the Social Returns to Innovation," in *Innovation and Public Policy*, ed. by B. F. Jones and A. Goolsbee, NBER.

Jones, C. (2022): "The Past and Future of Economic Growth: A Semi-Endogenous Perspective," *Annual Review of Economics.*

Jones, C. I. (1995): "R&D-Based Models of Economic Growth," *Journal of Political Economy.*

——— (2011): "Intermediate Goods and Weak Links in the Theory of Economic Development," *American Economic Journal: Macroeconomics*, 3, 1–28.

——— (2013): "Misallocation, Economic Growth, and Input-Output Economics," *Advances in Economics and Econometrics*, 2.

Jones, C. I. and J. Williams (1998): "Measuring the Social Return to R&D," *Quarterly Journal of Economics*.

Keller, W. (2004): "International Technology Diffusion," *Journal of Economic Literature*.

Kleinman, B., E. Liu, and S. Redding (2022): "International Friends and Enemies," *Working Paper*.

——— (forthcoming): "Dynamic Spatial General Equilibrium," *Econometrica*.

Koenig, M., M. Z. Song, K. Storesletten, and F. Zilibotti (forthcoming): "From Imitation to Innovation: Where Is All that Chinese R&D Going?" *Econometrica*.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological innovation, resource allocation, and growth," *Quarterly Journal of Economics*, 132, 665–712.

Lim, K. (2018): "Endogenous Production Networks and the Business Cycle," *Working Paper*.

Liu, E. (2019): "Industrial Policies in Production Networks*," *The Quarterly Journal of Economics*, 134, 1883–1948.

Liu, E. and A. Tsyvinski (2022): "Dynamical Structures and Spectral Properties of Input-Output Networks," *Working Paper*.

Lucking, B., N. Bloom, and J. Van Reenen (2018): "Have R&D Spillovers Changed?" *NBER Working Paper*.

Lybbert, T. J. and N. J. Zolas (2014): "Getting patents and economic data to speak to each other: An algorithmic links with probabilities approach for joint analyses of patenting and economic activity," *Research Policy*, 43, 530–542.

Ma, S. (2020): "The life cycle of corporate venture capital," *Review of Financial Studies*, 33, 358–394.

——— (2021): "Technological Obsolescence," *Working Paper*.

Machin, S. and J. Van Reenen (1998): "Technology and changes in skill structure: evidence from seven OECD countries," *The quarterly journal of economics*, 113, 1215–1244.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60, 531–542.

Melitz, M. and S. Redding (2021): "Trade and Innovation," *The Economics of Creative Destruction, Festschrift Volume in honor of Philippe Aghion and Peter Howitt*.

Ngai, L. R. and R. M. Samaniego (2011): "Accounting for research and productivity growth across industries," *Review of Economic Dynamics*.

Oberfield, E. (2018): "A Theory of Input-Output Architecture," *Econometrica*, 86.

Restuccia, D. and R. Rogerson (2008): "Policy distortions and aggregate productivity with heterogeneous establishment," *Review of Economic Dynamics*.

Santacreu, A. M. (2015): "Innovation, Diffusion, and Trade: Theory and Measurement," *Journal of Monetary Economics*.

Taschereau-Dumouchel, M. (2020): "Cascades and Fluctuations in an Economy with an Endogenous Production Network," *Working Paper*.

Thomson, R. (2017): "The effectiveness of R&D tax credits," *Review of Economics and Statistics*, 99, 544–549.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015): "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production," *Review of International Economics*.

vom Lehn, C. and T. Winberry (2022): "The Investment Network, Sectoral Comovement, and the Changing U.S. Business Cycle," *Quarterly Journal of Economics*.

Wilson, D. J. (2009): "Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits," *The Review of Economics and Statistics*, 91, 431–436.

Online Appendix (For Online Publication Only)

Table of Contents

A Proofs

A.1 Proof of Lemma 1: Optimal Labor Allocation

The planner's problem is

$$V^*\left(\{q_{i0}\}\right) \equiv \max_{\{\ell_{it}, s_{it}\}} \int_0^\infty e^{-\rho t} \sum_{i=1}^{K} \beta_i \ln y_{it} \, \mathrm{d}t,$$

subject to constraints (3), (4), (5), and (6). Substituting using (3) and (4), the objective can be re-written as

$$V^*\left(\{q_{i0}\}\right) \equiv \max_{\{\ell_{it}, s_{it}\}} \int_0^\infty e^{-\rho t} \sum_{i=1}^{K} \beta_i \ln q_{it}^{\psi} \ell_{it} \, \mathrm{d}t.$$

The FOC with regard to ℓ_{it} gives: $\frac{\beta_i}{\ell_{it}} = \frac{\beta_j}{\ell_{jt}}$. Therefore, for all t, $\ell_{it} = \beta_i \bar{\ell}$ for each sector i.

A.2 Proof of Proposition 1: Optimal R&D Allocation in the Baseline Model

The social planner's problem is

$$\begin{aligned} &\max_{\{\gamma_t\} \text{ s.t. } \gamma_t' 1 = 1 \forall t} \int_0^\infty e^{-\rho t} \boldsymbol{\beta}' \ln \boldsymbol{q}_t \, \mathrm{d}t \\ &\qquad \text{s.t. } \mathrm{d} \ln \boldsymbol{q}_t / \mathrm{d}t = \lambda \cdot \left(\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{\gamma}_t + \left(\boldsymbol{\Omega} - \boldsymbol{I} \right) \ln \boldsymbol{q}_t \right), \end{aligned}$$

The control variable is $\boldsymbol{\gamma}_t$ and the state variable is $\boldsymbol{q}_t$. Denote the co-state variables as $\boldsymbol{\mu}_t$. The current-value Hamiltonian writes

$$H(\gamma_t, \boldsymbol{q}_t, \boldsymbol{\mu}_t, \zeta) = \boldsymbol{\beta}' \ln \boldsymbol{q}_t + \lambda \boldsymbol{\mu}_t' \left(\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{\gamma}_t + \left(\boldsymbol{\Omega} - \boldsymbol{I} \right) \ln \boldsymbol{q}_t \right) + \zeta (1 - \boldsymbol{\gamma}_t' \mathbf{1}).$$

For notational simplicity we suppress dependence on time for the control, state, and co-state variables:

$$\begin{aligned} H(\{\gamma_i\}, \{q_i\}, \{\mu_i\}, \zeta, t) &= \sum_i \beta_i \ln q_i + \zeta (1 - \sum_i \gamma_i) \\ &\quad + \lambda \sum_i \mu_i \left(\ln \eta_i + \ln \bar{s} + \ln \gamma_i + \sum_j \omega_{ij} \ln q_j - \ln q_i \right) \end{aligned}$$

By the maximum principle

$$H_{\gamma_i} = 0 \Longleftrightarrow \frac{\lambda \mu_i}{\gamma_i} = \zeta \quad \forall i \tag{A1}$$

$$H_{\ln q_i} = \rho \mu_i - \dot{\mu}_i \Longleftrightarrow \beta_i - \lambda \mu_i + \lambda \sum_j \mu_j \omega_{ji} = \rho \mu_i - \dot{\mu}_i \tag{A2}$$

First, we show that the transversality condition $\lim_{t\to\infty} e^{-\rho t} H(\{\gamma_i\}, \{q_i\}, \{\mu_i\}, \zeta, t) = 0$ implies $\dot{\mu}_i = 0$ for all i. It is then immediate that the optimal R&D allocation $\boldsymbol{\gamma}$ is time invariant.

Note the matrix formula of equation (A2) is

$$\dot{\boldsymbol{\mu}}_t = [(\rho+\lambda)\boldsymbol{I} - \lambda\boldsymbol{\Omega}']\,\boldsymbol{\mu}_t - \boldsymbol{\beta} \tag{A3}$$

Then

$$\begin{aligned}\boldsymbol{\mu}_t &= e^{[(\rho+\lambda)\boldsymbol{I}-\lambda\boldsymbol{\Omega}']t}\boldsymbol{\mu}_0 - \left(\int_0^t e^{[(\rho+\lambda)\boldsymbol{I}-\lambda\boldsymbol{\Omega}'](t-s)}\,\mathrm{d}s\right)\boldsymbol{\beta} \\ &= e^{[(\rho+\lambda)\boldsymbol{I}-\lambda\boldsymbol{\Omega}']t}\boldsymbol{\mu}_0 - \left(e^{[(\rho+\lambda)\boldsymbol{I}-\lambda\boldsymbol{\Omega}']t} - \boldsymbol{I}\right)[(\rho+\lambda)\boldsymbol{I} - \lambda\boldsymbol{\Omega}']^{-1}\boldsymbol{\beta}.\end{aligned}$$

By transversality,

$$\begin{aligned}0 &= \lim_{t\to\infty} e^{-\rho t}\boldsymbol{\mu}_t \\ &= \lim_{t\to\infty} e^{[\lambda(\boldsymbol{I}-\boldsymbol{\Omega}')]t}\left[\boldsymbol{\mu}_0 - [(\rho+\lambda)\boldsymbol{I} - \lambda\boldsymbol{\Omega}']^{-1}\boldsymbol{\beta}\right].\end{aligned}$$

Hence it must be the case that $\boldsymbol{\mu}_0 = [(\rho+\lambda)\boldsymbol{I} - \lambda\boldsymbol{\Omega}']^{-1}\boldsymbol{\beta}$. Plugging it to the explicit solution of $\boldsymbol{\mu}_t$ and then back to (A3), we can get $\dot{\boldsymbol{\mu}}_t = 0$. Hence $\boldsymbol{\mu}_t$ and $\boldsymbol{\gamma}_t$ are time invariant.

We then can calculate $\boldsymbol{\gamma}$. First obtain $\boldsymbol{\mu}$ directly from FOC (A3):

$$(\rho+\lambda)\boldsymbol{\mu}_t'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right) = \boldsymbol{\beta}' \iff \boldsymbol{\mu}_t' = \frac{1}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}.$$

According to Equation (A1), $\boldsymbol{\gamma}$ is proportional to $\boldsymbol{\mu}$ and subject to $\sum_i \gamma_i = 1$. We can then find $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1},$$

since

$$\begin{aligned}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}\boldsymbol{1} &= \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\sum_{s=0}^{\infty}\left(\frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{s}\boldsymbol{1}\right) \\ &= \frac{\rho}{\rho+\lambda}\sum_{s=0}^{\infty}\left(\frac{1}{1+\rho/\lambda}\right)^{s} \\ &= 1,\end{aligned}$$

as desired.

A.3 Proof of Lemma 2: Economic Growth Rate Along a Balanced Growth Path

Consider a BGP in which R&D allocation shares follow the vector $\boldsymbol{b}$ and the growth rate of sectoral knowledge stock is time-invariant. The law of motion for stock vector is

$$\mathrm{d}\ln \boldsymbol{q}_t/\mathrm{d}t = \lambda \cdot (\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t).$$

Taking derivative with respect to time,

$$\boldsymbol{0} = \lambda (\boldsymbol{\Omega} - \boldsymbol{I}) \frac{\mathrm{d}\ln \boldsymbol{q}_t}{\mathrm{d}t},$$

implying that the vector of sectoral growth rates $\frac{\mathrm{d}\ln \boldsymbol{q}_t}{\mathrm{d}t}$ is the right-Perron eigenvector of $\boldsymbol{\Omega}$. Because $\boldsymbol{\Omega}$ is a row-stochastic matrix, this implies that $\frac{\mathrm{d}\ln \boldsymbol{q}_t}{\mathrm{d}t}$ must be a constant vector, meaning the knowledge stock in every sector must grow at the same rate $g^q(\boldsymbol{b})$. Hence,

$$g^q(\boldsymbol{b})\boldsymbol{1} = \frac{\mathrm{d}\ln \boldsymbol{q}_t}{\mathrm{d}t} = \lambda \cdot (\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t). \tag{A4}$$

Left-multiply by the centrality $\boldsymbol{a}'$ of $\boldsymbol{\Omega}$ on both sides:

$$\begin{aligned} g^q(\boldsymbol{b}) &= \boldsymbol{a}' \cdot g(\boldsymbol{b}) \boldsymbol{1} \\ &= \lambda \cdot (\boldsymbol{a}' \ln \boldsymbol{\eta} + \boldsymbol{a}' \cdot \boldsymbol{1} \ln \bar{s} + \boldsymbol{a}' \ln \boldsymbol{b} + \boldsymbol{a}'(\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t) \\ &= \lambda \cdot (\boldsymbol{a}' \ln \boldsymbol{\eta} + \ln \bar{s} + \boldsymbol{a}' \ln \boldsymbol{b}) \\ &= \text{const} + \lambda \cdot \boldsymbol{a}' \ln \boldsymbol{b}. \end{aligned}$$

The third equation is based on the properties of the innovation centrality vector: $\boldsymbol{a}' = \boldsymbol{a}'\boldsymbol{\Omega}$ and $\sum_{i=1}^{K} a_i = 1$. That $g^y(\boldsymbol{b}) = \psi \cdot g^q(\boldsymbol{b})$ is immediate from the production function $y_i = q_i^{\psi} \ell_i$.

A.4 Proof of Proposition 2

Starting from $\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$, right-multiply both sides by $\frac{\rho+\lambda}{\lambda}\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)$ to get

$$\boldsymbol{\gamma}' \left(\frac{\rho+\lambda}{\lambda}\boldsymbol{I} - \boldsymbol{\Omega}\right) = \frac{\rho}{\lambda}\boldsymbol{\beta}' \quad \Longleftrightarrow \quad \boldsymbol{\gamma}'(\boldsymbol{I} - \boldsymbol{\Omega}) + \frac{\rho}{\lambda}(\boldsymbol{\gamma}' - \boldsymbol{\beta}') = \boldsymbol{0}'.$$

Taking the limit as $\rho/\lambda \to 0$, $\boldsymbol{\gamma}'(\boldsymbol{I} - \boldsymbol{\Omega}) \to \boldsymbol{0}$ implies $\boldsymbol{\gamma} \to \boldsymbol{a}$; taking the limit as $\rho/\lambda \to \infty$, $\boldsymbol{\gamma} \to \boldsymbol{\beta}$, as desired.

A.5 Proof of Proposition 3: Welfare Impact of R&D Reallocation

The law of motion for knowledge stock $\ln \boldsymbol{q}$ under R&D allocation $\boldsymbol{b}$ is

$$\frac{\mathrm{d}\ln \boldsymbol{q}}{\mathrm{d}t} = \lambda (\ln \boldsymbol{\eta} + \ln \bar{s} \cdot \boldsymbol{1} + \ln \boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q})$$

Let $\boldsymbol{a}$ denote the left-eigenvector centrality of $\boldsymbol{\Omega}$ (normalized to sum to one). We separately analyze $\boldsymbol{a}'\ln\boldsymbol{q}_t$, i.e., the centrality-weighted average knowledge stock, and the deviation of knowledge stock from this average, $(\boldsymbol{I}-\boldsymbol{1a}')\ln\boldsymbol{q}_t$.[9] We first show the former always grows at a constant rate even away from a BGP, whereas the latter converges to a constant vector as the economy converges to a BGP.

From the law of motion, we know

$$\begin{aligned}\boldsymbol{a}'\frac{\mathrm{d}\ln\boldsymbol{q}}{\mathrm{d}t} &= \lambda\left(\boldsymbol{a}'\ln\boldsymbol{\eta}+\ln\bar{s}'\cdot\boldsymbol{a}'\boldsymbol{1}+\boldsymbol{a}'\ln\boldsymbol{b}+\boldsymbol{a}'\left(\boldsymbol{\Omega}-\boldsymbol{I}\right)\ln\boldsymbol{q}\right)\\ &= \lambda\boldsymbol{a}'\left(\ln\boldsymbol{\eta}+\ln\bar{s}\cdot\boldsymbol{1}+\ln\boldsymbol{b}\right)\end{aligned}$$

Hence, given time-invariant R&D allocation $\boldsymbol{b}$, $\boldsymbol{a}'\ln\boldsymbol{q}_t$ always grows at a constant rate (and it equals to the rate of growth along a BGP) and can be solved in closed-form:

$$\boldsymbol{a}'\ln\boldsymbol{q}_t=\boldsymbol{a}'\ln\boldsymbol{q}_0+\lambda\boldsymbol{a}'\left(\ln\boldsymbol{\eta}+\ln\bar{s}\cdot\boldsymbol{1}+\ln\boldsymbol{b}\right)t$$

Note that $\boldsymbol{a}'\boldsymbol{1}=1$; hence $(\boldsymbol{I}-\boldsymbol{1a}')\left(\ln\bar{s}\cdot\boldsymbol{1}\right)=\boldsymbol{0}$. Let $\boldsymbol{A}\equiv\boldsymbol{1a}'$. Note that the row-stochastic matrix $\boldsymbol{\Omega}$ represents a Markov chain, for which $\boldsymbol{a}$ is the stationary distribution, and $\boldsymbol{A}\equiv\lim_{s\to\infty}\boldsymbol{\Omega}^s$. Also note that

$$\begin{aligned}\left(\boldsymbol{I}-\boldsymbol{A}\right)\left(\boldsymbol{\Omega}-\boldsymbol{I}\right) &= \left(\boldsymbol{\Omega}-\boldsymbol{I}\right)\left(\boldsymbol{I}-\boldsymbol{A}\right)\\ &= -\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)\left(\boldsymbol{I}-\boldsymbol{A}\right)\end{aligned}$$

Left-multiply the law of motion by $(\boldsymbol{I}-\boldsymbol{A})$, substitute the above, and let $\widetilde{\ln\boldsymbol{q}_t}\equiv(\boldsymbol{I}-\boldsymbol{A})\ln\boldsymbol{q}_t$, we get

$$\frac{\mathrm{d}\widetilde{\ln\boldsymbol{q}_t}}{\mathrm{d}t} = \lambda\left(\boldsymbol{I}-\boldsymbol{A}\right)\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)-\lambda\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)\widetilde{\ln\boldsymbol{q}_t}$$

We can integrate the ODE system:

$$\begin{aligned}\widetilde{\ln\boldsymbol{q}_t} &= e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\left[\widetilde{\ln\boldsymbol{q}_0}+\lambda\int_0^t e^{\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})s}\left(\boldsymbol{I}-\boldsymbol{A}\right)\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)\,\mathrm{d}s\right]\\ &= e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\widetilde{\ln\boldsymbol{q}_0}+\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left(\boldsymbol{I}-e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right)\left(\boldsymbol{I}-\boldsymbol{A}\right)\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)\end{aligned}$$

Which implies that there's a closed-form solution for the sectoral knowledge stock along the entire path of the economy:

$$\begin{aligned}\ln\boldsymbol{q}_t &= \widetilde{\ln\boldsymbol{q}_t}+\boldsymbol{A}\ln\boldsymbol{q}_t\\ &= \boldsymbol{A}\ln\boldsymbol{q}_0+\lambda\boldsymbol{A}\left(\ln\boldsymbol{\eta}+\ln\bar{s}\cdot\boldsymbol{1}+\ln\boldsymbol{b}\right)t\\ &\quad +e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\widetilde{\ln\boldsymbol{q}_0}+\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left(\boldsymbol{I}-e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right)\left(\boldsymbol{I}-\boldsymbol{A}\right)\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)\end{aligned}$$

Starting from the same initial knowledge stock $\boldsymbol{q}_0$ but with two different time-invariant R&D

[9]We separate these two objects because, the matrix $(\boldsymbol{I}-\boldsymbol{\Omega})$ is not invertible, but $(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{1a}')$ generically is. The proof shown below utilizes the invertibility of $(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{1a}')$ to solve for $(\boldsymbol{I}-\boldsymbol{1a}')\ln\boldsymbol{q}_t$.

allocations $\widetilde{\boldsymbol{b}}$ and $\boldsymbol{b}$, we have the following difference in knowledge stock over time:

$$\begin{aligned}
\ln \boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right) - \ln \boldsymbol{q}_t\left(\boldsymbol{b}\right) &= \boldsymbol{A} \ln \boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right) - \boldsymbol{A} \ln \boldsymbol{q}_t\left(\boldsymbol{b}\right) \\
&\quad + \widetilde{\ln \boldsymbol{q}_t}\left(\widetilde{\boldsymbol{b}}\right) - \widetilde{\ln \boldsymbol{q}_t}\left(\boldsymbol{b}\right) \\
&= \left[\boldsymbol{A}\lambda t + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)^{-1}\left(\boldsymbol{I} - e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right)\left(\boldsymbol{I} - \boldsymbol{A}\right)\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right)
\end{aligned}$$

Note

$$\int_0^\infty e^{-\rho t}\lambda t \,\mathrm{d}t = -\frac{1}{\rho} e^{-\rho t}\lambda t\Big|_0^\infty + \int_0^\infty \frac{1}{\rho} e^{-\rho t}\lambda \,\mathrm{d}t = \frac{\lambda}{\rho^2}$$

The difference in consumer welfare under two time-invariant paths of R&D allocations is

$$\begin{aligned}
& V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \widetilde{\boldsymbol{b}}\right) - V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \boldsymbol{b}\right) \\
= \; & \psi \boldsymbol{\beta}' \int_0^\infty e^{-\rho t}\left[\ln \boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right) - \ln \boldsymbol{q}_t\left(\boldsymbol{b}\right)\right] \mathrm{d}t \\
= \; & \psi \boldsymbol{\beta}' \int_0^\infty e^{-\rho t}\left[\boldsymbol{A}\lambda t + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)^{-1}\left(\boldsymbol{I} - e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right)\left(\boldsymbol{I} - \boldsymbol{A}\right)\right] \mathrm{d}t \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \frac{\psi\lambda}{\rho^2}\boldsymbol{\beta}'\boldsymbol{A}\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
& + \psi \boldsymbol{\beta}'\left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)^{-1}\left[\frac{1}{\rho}\boldsymbol{I} - \int_0^\infty \left(e^{-((\rho+\lambda)\boldsymbol{I} - \lambda(\boldsymbol{\Omega}-\boldsymbol{A}))t}\right) \mathrm{d}t\right]\left(\boldsymbol{I} - \boldsymbol{A}\right)\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \psi \boldsymbol{\beta}'\left[\frac{\lambda}{\rho^2}\boldsymbol{A} + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)^{-1}\left[\frac{1}{\rho}\boldsymbol{I} - \frac{1}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)^{-1}\right]\left(\boldsymbol{I} - \boldsymbol{A}\right)\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \psi \boldsymbol{\beta}'\left[\frac{\lambda}{\rho^2}\boldsymbol{A} + \frac{1}{\rho}\left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)^{-1}\left[\frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I} - \left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)^{-1}\right]\left(\boldsymbol{I} - \boldsymbol{A}\right)\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \frac{\psi}{\rho}\boldsymbol{\beta}'\left[\frac{\lambda}{\rho}\boldsymbol{A} + \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)^{-1}\left(\boldsymbol{I} - \boldsymbol{A}\right)\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \frac{\psi}{\rho}\boldsymbol{\beta}'\frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)^{-1}\left[\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)\frac{\rho+\lambda}{\rho}\boldsymbol{A} + \left(\boldsymbol{I} - \boldsymbol{A}\right)\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \\
= \; & \frac{\psi}{\rho^2}\boldsymbol{\beta}'\frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{\Omega} - \boldsymbol{A}\right)\right)^{-1}\left[\rho\boldsymbol{I} + \lambda\boldsymbol{A}\right]\left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right)
\end{aligned}$$

Note

$$\left(\rho\boldsymbol{I} + \lambda\boldsymbol{A}\right)^{-1} = \frac{1}{\rho}\left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{A}\right)$$

To see this,

$$
\begin{aligned}
& (\rho \boldsymbol{I}+\lambda \boldsymbol{A}) \frac{1}{\rho}\left(\boldsymbol{I}-\frac{1}{1+\rho / \lambda} \boldsymbol{A}\right) \\
= & \frac{1}{\rho}\left(\rho \boldsymbol{I}+\lambda \boldsymbol{A}-(\rho \boldsymbol{I}+\lambda \boldsymbol{A}) \frac{1}{1+\rho / \lambda} \boldsymbol{A}\right) \\
= & \boldsymbol{I}+\frac{1}{\rho}(\lambda \boldsymbol{A}-\lambda \boldsymbol{A}) \\
= & \boldsymbol{I}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& V\left(\boldsymbol{q}_0 ;\left\{\boldsymbol{\ell}_t\right\}, \widetilde{\boldsymbol{b}}\right)-V\left(\boldsymbol{q}_0 ;\left\{\boldsymbol{\ell}_t\right\}, \boldsymbol{b}\right) \\
= & \frac{\psi}{\rho} \boldsymbol{\beta}^{\prime} \frac{\lambda}{\rho+\lambda}\left(\left(\boldsymbol{I}-\frac{1}{1+\rho / \lambda} \boldsymbol{A}\right)\left[\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}(\boldsymbol{\Omega}-\boldsymbol{A})\right]\right)^{-1}\left(\ln \widetilde{\boldsymbol{b}}-\ln \boldsymbol{b}\right) \\
= & \frac{\psi}{\rho} \boldsymbol{\beta}^{\prime} \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}(\boldsymbol{\Omega}-\boldsymbol{A})-\frac{1}{1+\rho / \lambda} \boldsymbol{A}\right)^{-1}\left(\ln \widetilde{\boldsymbol{b}}-\ln \boldsymbol{b}\right) \\
= & \frac{\psi}{\rho} \boldsymbol{\beta}^{\prime} \frac{\lambda}{\rho+\lambda}\left(\boldsymbol{I}-\frac{1}{1+\rho / \lambda} \boldsymbol{\Omega}\right)^{-1}\left(\ln \widetilde{\boldsymbol{b}}-\ln \boldsymbol{b}\right) \\
= & \frac{\psi \lambda}{\rho^2} \boldsymbol{\gamma}^{\prime}\left(\ln \widetilde{\boldsymbol{b}}-\ln \boldsymbol{b}\right),
\end{aligned}
$$

as desired.

A.6 Proof of Proposition 4: Consumption-Equivalent Welfare Gains from Adopting the Optimal R&D

For a given consumption path $\{y_t\}$, the welfare gain under the alternative consumption path $\{\mathcal{L} \cdot y_t\}$ is $\int e^{-\rho t} \ln \mathcal{L} \, \mathrm{d}t = \frac{\ln \mathcal{L}}{\rho}$. The result thus immediately follows Proposition 3.

A.7 Proof of Proposition 5: General Functional Forms and Endogenous Innovation Network

Consider the economic environment outlined in Section 2.5, with preferences

$$
\int_0^{\infty} e^{-\rho t} \ln \mathcal{Y}\left(\left\{q_{it}^{\psi} \ell_{it}\right\}\right) \mathrm{d}t
$$

and knowledge stock law of motion

$$
\mathrm{d} \ln q_{it} / \mathrm{d}t = f\left(\ln \left(b_{it} \bar{s}\right)+\ln \mathcal{X}_i\left(\left\{q_{jt}\right\}\right)\right) \quad \forall i,
$$

where ℓ_{it} is the measure of production workers allocated to each variety in sector i at time t.

Consider the economy initially at $t = 0$ in a BGP with R&D allocation $\boldsymbol{b}$. Define

$$\beta_i \equiv \frac{\partial \ln \mathcal{Y}\left(\{y_{jt}\}\right)}{\partial \ln y_{it}}\bigg|_{t=0}, \qquad \omega_{ij} \equiv \begin{cases} \frac{\partial \ln \mathcal{X}_i(\{q_{kt}\})}{\partial \ln q_{jt}}\Big|_{t=0} & \text{if } i \neq j \\ 1 + \frac{\partial \ln \mathcal{X}_i(\{q_{it}\})}{\partial \ln q_{it}}\Big|_{t=0} & \text{otherwise.} \end{cases}$$

$\boldsymbol{\beta} \equiv [\beta_i]$ and $\boldsymbol{\Omega} \equiv [\omega_{ij}]$ are the consumption and innovation spillover elasticities evaluated in the initial BGP. Note that (1) $\mathcal{X}_i(\cdot)$ being homogeneous-of-degree-zero with positive cross-sector spillovers and (2) $|\partial \ln \mathcal{X}_i(\cdot) / \partial \ln q_{jt}| \leq 1 \forall i, j$ jointly imply that $\omega_{ij} \geq 0$ for all i, j. Let $\lambda \equiv f'(\cdot)$ denote the slope of the function f, and define $\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$.

We now derive the first-order welfare impact of perturbing R&D allocation. Let $V(\ln \boldsymbol{q}_0; \ln \boldsymbol{b})$ denote the welfare under log-R&D allocation $\ln \boldsymbol{b}$. Formally, we show that the Gateaux derivative of welfare with respect to log R&D allocation $\ln \boldsymbol{b}$ in the direction of $\boldsymbol{h}$ is

$$\lim_{\alpha \to 0} \frac{V(\ln \boldsymbol{q}_0; \ln \boldsymbol{b} + \alpha \boldsymbol{h}) - V(\ln \boldsymbol{q}_0; \boldsymbol{b})}{\alpha} = \frac{\psi \lambda}{\rho^2} \boldsymbol{\gamma}' \boldsymbol{h}.$$

Given log-R&D allocation $\ln \boldsymbol{b} + \alpha \boldsymbol{h}$, the law of motion for knowledge stock satisfies

$$\frac{\mathrm{d} \ln \boldsymbol{q}_t}{\mathrm{d}t} = f\left(\ln \boldsymbol{b} + \alpha \boldsymbol{h} + \ln \boldsymbol{\chi}\left(\{\ln \boldsymbol{q}_t\}\right)\right)$$

$$\frac{\partial^2 \ln \boldsymbol{q}_t}{\partial \alpha \partial t} = \lambda \boldsymbol{h} + \lambda\left(\boldsymbol{\Omega} - \boldsymbol{I}\right) \frac{\partial \ln \boldsymbol{q}_t}{\partial \alpha}$$

$$\implies \frac{\partial \ln \boldsymbol{q}_t}{\partial \alpha} = \left(\boldsymbol{I} - \boldsymbol{\Omega}\right)^{-1}\left[\boldsymbol{I} - e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega})t}\right]\boldsymbol{h}$$

$$
\begin{aligned}
&\lim_{\alpha\to 0}\frac{V\left(\ln \boldsymbol{q}_0;\ln \boldsymbol{b}+\alpha \boldsymbol{h}\right)-V\left(\ln \boldsymbol{q}_0;\boldsymbol{b}\right)}{\alpha}\\
= & \int e^{-\rho t}\frac{\partial \ln Y\left(\psi \ln \boldsymbol{q}_t\right)}{\partial \ln \boldsymbol{q}_t}\frac{\partial \ln \boldsymbol{q}_t}{\partial \alpha}\,\mathrm{d}t\\
= & \psi \boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\int e^{-\rho t}\left[\boldsymbol{I}-e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega})t}\right]\,\mathrm{d}t\boldsymbol{h}\\
= & \psi \boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\left[\frac{1}{\rho}\boldsymbol{I}-\int e^{-((\rho+\lambda)\boldsymbol{I}-\lambda\boldsymbol{\Omega})t}\,\mathrm{d}t\right]\boldsymbol{h}\\
= & \psi \boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\left[\frac{1}{\rho}\boldsymbol{I}-\frac{1}{\rho+\lambda}\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right)^{-1}\right]\boldsymbol{h}\\
= & \psi \boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\left[\frac{1}{\rho}\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right)-\frac{1}{\rho+\lambda}\boldsymbol{I}\right]\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right)^{-1}\boldsymbol{h}\\
= & \frac{\psi}{\rho}\boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\left[\frac{\lambda}{\rho+\lambda}\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right]\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right)^{-1}\boldsymbol{h}\\
= & \frac{\psi\lambda}{\rho^2}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I}-\frac{\lambda}{\rho+\lambda}\boldsymbol{\Omega}\right)^{-1}\boldsymbol{h}\\
= & \frac{\psi\lambda}{\rho^2}\boldsymbol{\gamma}'\boldsymbol{h},
\end{aligned}
$$

as desired.

A.8 Proof of Proposition 6: Optimal R&D in the Presence of Foreign Spillovers

First, note that given output y_t and the price of imports p_t^f, consumption, export, and import must solve

$$
\bar{\mathcal{C}}^*\left(y_t,p_t^f\right)\equiv\max_{c_t^d,c_t^f}\mathcal{C}\left(c_t^d,c_t^f\right)\quad \text{s.t. } y_t-c_t^d=p_t^f c_t^f. \tag{A5}
$$

Since $\mathcal{C}\left(\cdot\right)$ features constant-returns-to-scale, we can re-write the maximized consumption aggregator as $\bar{\mathcal{C}}^*\left(y_t,p_t^f\right)=y_t\mathcal{C}^*\left(p_t^f\right)$ for some function $\mathcal{C}^*$. Hence, for any $\boldsymbol{q}_t$, $\{\ell_{it}\}$ are chosen to maximize flow output; thus the optimal worker allocation features $\ell_{it}/\bar{\ell}=\beta_i$ as in the closed economy.

We next characterize the optimal R&D allocation. Let $\boldsymbol{\Theta}\equiv\boldsymbol{\Omega}\circ\boldsymbol{X}$. Given the law of motion for sectoral knowledge stock, we can solve for the evolution of knowledge stock in closed form as a function of R&D allocation $\boldsymbol{b}_t$:

$$
\ln \boldsymbol{q}_t=e^{\lambda(\boldsymbol{\Theta}-\boldsymbol{I})t}\left[\ln \boldsymbol{q}_0+\lambda\int_0^t e^{-\lambda(\boldsymbol{\Theta}-\boldsymbol{I})s}\left(\left(\boldsymbol{\Omega}-\boldsymbol{\Theta}\right)\ln \boldsymbol{q}_s^f+\ln\boldsymbol{\eta}+\ln\bar{s}+\ln\boldsymbol{b}_s\right)\,\mathrm{d}s\right]. \tag{A6}
$$

The optimal R&D allocation is

$$\begin{aligned}\{\boldsymbol{\gamma}_t\} &= \arg\max_{\{\boldsymbol{b}_s\}} \int_0^\infty e^{-\rho t} \ln \bar{\mathcal{C}}^* \left(y_t(\{\boldsymbol{b}_s\}), p_t^f \right) \mathrm{d}t \\ &= \arg\max_{\{\boldsymbol{b}_s\}} \int_0^\infty e^{-\rho t} \ln y_t(\{\boldsymbol{b}_s\}) \, \mathrm{d}t \\ &= \arg\max_{\{\boldsymbol{b}_s\}} \int_0^\infty e^{-\rho t} \boldsymbol{\beta}' \ln \boldsymbol{q}_t(\{\boldsymbol{b}_s\}) \, \mathrm{d}t \\ &= \arg\max_{\{\boldsymbol{b}_s\}} \boldsymbol{\beta}' \int_0^\infty e^{-\rho t} \left[\lambda \int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Theta})(t-s)} \ln \boldsymbol{b}_s \, \mathrm{d}s \right] \mathrm{d}t.\end{aligned}$$

The optimal R&D allocation therefore coincides with the solution to the following problem:

$$\arg\max_{\{\boldsymbol{b}_s\}} \int_0^\infty e^{-\rho t} \boldsymbol{\beta}' \boldsymbol{m}_t \, \mathrm{d}t$$

$$\text{s.t. } \dot{\boldsymbol{m}}_t = \lambda (\boldsymbol{\Theta} - \boldsymbol{I}) \boldsymbol{m}_t + \lambda \ln \boldsymbol{b}_t, \quad \boldsymbol{m}_0 \text{ given},$$

which can be solved in closed form by forming the Hamiltonian, following a similar procedure as in the proof for Proposition 1. The solution features

$$\boldsymbol{\gamma}' = \xi^{-1} \frac{\rho}{\rho+\lambda} \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega} \circ \boldsymbol{X}}{1+\rho/\lambda} \right)^{-1}, \quad \xi \equiv \frac{\rho}{\rho+\lambda} \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega} \circ \boldsymbol{X}}{1+\rho/\lambda} \right)^{-1} \boldsymbol{1},$$

as desired.

A.9 Proof of Proposition 7: Welfare Impact of R&D in the Presence of Foreign Spillovers

Starting from an initial condition $\boldsymbol{q}_0$, a path of foreign knowledge and import prices $\left\{ \boldsymbol{q}_t^f, p_t^f \right\}$, and a path of worker allocation $\{\boldsymbol{\ell}_t\}$, the welfare differences between an economy with optimal R&D allocation $\boldsymbol{\gamma}$ and an economy with time-invariant allocation $\boldsymbol{b}$ is

$$V(\boldsymbol{\gamma}) - V(\boldsymbol{b}) = \int_0^\infty e^{-\rho t} \left[\ln \bar{\mathcal{C}}^* \left(y_t(\boldsymbol{\gamma}), p_t^f \right) - \ln \bar{\mathcal{C}}^* \left(y_t(\boldsymbol{b}), p_t^f \right) \right] \mathrm{d}t,$$

where $\bar{\mathcal{C}}^*$ is defined in (A5). Following the proof to Proposition 6, $\bar{\mathcal{C}}^* \left(y_t, p_t^f \right) = y_t \mathcal{C}^* \left(p_t^f \right)$; hence the welfare differences can be re-written as

$$V(\boldsymbol{\gamma}) - V(\boldsymbol{b}) = \int_0^\infty e^{-\rho t} \left[\ln y_t(\boldsymbol{\gamma}) - \ln y_t(\boldsymbol{b}) \right] \mathrm{d}t.$$

Since $\ln y_t$ is additive in $\psi \boldsymbol{\beta}' \ln \boldsymbol{q}_t$, we can re-write the welfare differences in terms of the discounted integral of $\boldsymbol{\beta}$-weighted differences in knowledge stock induced by the two different R&D

allocation vectors. By (A6), we can re-write the welfare differences as

$$V(\gamma) - V(\boldsymbol{b}) = \psi\boldsymbol{\beta}' \int_0^\infty e^{-\rho t} \left[\lambda \int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Theta})(t-s)}\, \mathrm{d}s\right] \mathrm{d}t \left(\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}\right),$$

where $\boldsymbol{\Theta} \equiv \boldsymbol{\Omega} \circ \boldsymbol{X}$. To simplify the integral we follow the proof to Proposition 3:[10]

$$\begin{aligned}
V(\gamma) - V(\boldsymbol{b}) &= \psi\boldsymbol{\beta}' \int_0^\infty e^{-\rho t} \left[\lambda \int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Theta})(t-s)}\, \mathrm{d}s\right] \mathrm{d}t \left(\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}\right), \\
&= \psi\boldsymbol{\beta}' (\boldsymbol{I}-\boldsymbol{\Theta})^{-1} \left(\int_0^\infty e^{-\rho t} \left[\boldsymbol{I} - e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Theta})t}\right] \mathrm{d}t\right) (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}) \\
&= \frac{\psi}{\rho}\boldsymbol{\beta}' (\boldsymbol{I}-\boldsymbol{\Theta})^{-1} \left(\boldsymbol{I} - \frac{\rho}{\rho+\lambda}\left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{\Theta}\right)^{-1}\right) (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}) \\
&= \frac{\psi\lambda}{\rho^2}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{\Theta}\right)^{-1} (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}) \\
&= \frac{\psi\lambda}{\rho^2}\underbrace{\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{\Theta}\right)^{-1}\boldsymbol{1}}_{\equiv\xi}\underbrace{\frac{\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{\Theta}\right)^{-1}}{\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{1}{1+\rho/\lambda}\boldsymbol{\Theta}\right)^{-1}\boldsymbol{1}}}_{\equiv\boldsymbol{\gamma}'} (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}) \\
&= \frac{\psi\lambda}{\rho^2}\xi\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}).
\end{aligned}$$

For a given consumption path $\left\{\bar{\mathcal{C}}^*\left(y_t, p_t^f\right)\right\}$, the welfare gain under the alternative consumption path $\left\{\mathcal{L}\cdot\bar{\mathcal{C}}^*\left(y_t, p_t^f\right)\right\}$ is $\int e^{-\rho t} \ln \mathcal{L}\, \mathrm{d}t = \frac{\ln \mathcal{L}}{\rho}$. The consumption-equivalent welfare gains from adopting the optimal R&D allocation is thus

$$\mathcal{L}(\boldsymbol{b}, \xi) = \exp\left(\frac{\psi\lambda}{\rho}\xi\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})\right),$$

as desired.

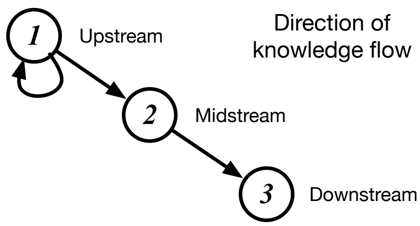
B Theoretical Extensions

B.1 Three-Sector Example

To demonstrate Propositions 1 and 2, consider the following three-sector example, where knowledge flows from sector 1 to sector 2 and from sector 2 to sector 3. Sector 1 can thus be interpreted as the "upstream" sector of knowledge flows, and sector 3 is the knowledge "downstream." To ensure the knowledge aggregator χ_{it} has constant returns to scale in every sector, we specify that

[10]Note that $\boldsymbol{I} - \boldsymbol{\Theta}$ is generically invertible—the economy with foreign spillovers exhibit aggregate decreasing-returns-to-scale in domestic R&D—so the proof here is simpler than in the baseline model.

knowledge in sector 1 also benefits itself. For simplicity, we assume the consumer values goods from each sector equally, with consumption share $\beta_i = 1/3$ for all i.



$$\boldsymbol{\Omega} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

The socially optimal R&D allocations depend on the effective discount rate ρ/λ and should follow, according to Proposition 1,

$$\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1} = \left[\begin{array}{ccc} \frac{1+(1+\rho/\lambda)+(1+\rho/\lambda)^2}{3(1+\rho/\lambda)^2} & \frac{\rho/\lambda+\rho/\lambda(1+\rho/\lambda)}{3(1+\rho/\lambda)^2} & \frac{\rho/\lambda}{3(1+\rho/\lambda)} \end{array}\right].$$

When the effective discount rate ρ/λ is lower, more resources should be directed to upstream sector 1 and fewer to downstream sector 3. A myopic planner ($\rho/\lambda \to \infty$) chooses $\gamma_1 = \gamma_3$; as the when $\rho/\lambda = 1$, $\gamma_1/\gamma_3 \approx 3.5$; when $\rho/\lambda = 0.1$, $\gamma_1/\gamma_3 \approx 30.1$.

B.2 Embedding Input-Output Linkages into Production Functions

We now expand on Section 2.7.1 and introduce input-output linkages into the baseline model. As discussed in the main text, for the optimal R&D allocation $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$, the presence of a production network requires a different construction for the $\boldsymbol{\beta}$ vector, but the innovation network $\boldsymbol{\Omega}$ term is unaffected. Formally, the $\boldsymbol{\beta}$ vector should capture the elasticity of aggregate consumption with respect to the knowledge stock in each sector; in the presence of a production network, it should reflect not only the consumer preferences but also the production network structure. With this adjustment, our main results continue to hold in this environment.

Specifically, suppose the production of good i requires other goods as intermediate inputs:

$$\ln y_{it} = \textstyle\sum_{j=1}^{K}\sigma_{ij}\ln m_{ijt} + \alpha_i \ln q_{it}^{\psi}\ell_{it}\,\mathrm{d}\nu, \quad \alpha_i + \textstyle\sum_{j=1}^{K}\sigma_{ij} = 1, \tag{A7}$$

where m_{ijt} is the quantity of good j used for the production of good i, α_i is sector i's output elasticity to value-added, and σ_{ij} is sector i's output elasticity to input j. The baseline model is a special case with $\sigma_{ij} = 0$ for all i, j. When an equal amount of labor ℓ_{it} is allocated to each variety within a sector, production function (A7) takes the standard form in the canonical production network model (Acemoglu et al., 2012):

$$y_{it} = \left(q_{it}^{\psi}\ell_{it}\right)^{\alpha_i}\textstyle\prod_{j=1}^{K} m_{ijt}^{\sigma_{ij}}. \tag{A8}$$

The market clearing condition for sectoral good follows

$$y_{jt} = \sum_i m_{ijt} + c_{jt}. \tag{A9}$$

The aggregate consumption bundle follows:

$$\ln y_t = \sum_{i=1}^{K} \beta_i \ln c_{it}. \tag{A10}$$

Consider the problem of choosing worker allocation to maximize flow consumption:

$$\ln y^* (\boldsymbol{q}_t) \equiv \max_{\{\ell_{it}\}} \sum_{i=1}^{K} \beta_i \ln c_{it}$$

subject to (A9) and (A8). Let $\boldsymbol{\Sigma} \equiv [\sigma_{ij}]$ denote the matrix of input-output elasticities. Standard results in the production networks literature (e.g., see Acemoglu et al., 2012 and Liu, 2019) imply

$$\ln y^* (\boldsymbol{q}_t) = \text{const} + \ln \bar{\ell} + \sum_i \hat{\beta}_i \ln q_{it},$$

where $\hat{\beta}_i \equiv \alpha_i \left[\boldsymbol{\beta}' (\boldsymbol{I} - \boldsymbol{\Sigma})^{-1}\right]_i$ is the product between sectoral value-added elasticity α_i and the i-th entry of the influence vector $\boldsymbol{\beta}' (\boldsymbol{I} - \boldsymbol{\Sigma})^{-1}$. $\hat{\beta}_i$ can be interpreted as the elasticity of aggregate output with respect to sectoral knowledge stock. Hence, results in the main text extend intuitively to this setting with input-output linkages: the optimal worker allocation follows the vector $\hat{\boldsymbol{\beta}}$, and the optimal R&D allocation $\gamma_{it} \equiv s_{it}/\bar{s}$ follows $\boldsymbol{\gamma}' \propto \hat{\boldsymbol{\beta}}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$.

B.3 Semi-Endogenous Growth

Our baseline model features endogenous growth: a positive growth rate of aggregate output along a balanced growth path in the absence of population growth. This is because the R&D technology features aggregate constant-returns-to-scale in sectoral knowledge stock. We now expand on Section 2.7.2 and embed our innovation network formulation into a semi-endogenous growth setting, with a constant growth rate in the total measure of scientists $\bar{s}_t = \bar{s}_0 e^{\bar{g}t}$. We show that the optimal R&D allocation follows $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}$, and the consumption-equivalent welfare impact of adopting the optimal allocation is $\mathcal{L}(\boldsymbol{b}) = \exp\left(\frac{\lambda}{\rho+\kappa\lambda}\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})\right)$.

Specifically, replace the knowledge stock evolution equation (5) with

$$\dot{q}_{it}/q_{it} = \lambda \ln \left(n_{it}/q_{it}^{1+\kappa}\right),$$

where $\kappa \geq 0$ captures the rate at which proportional improvements in knowledge are getting harder to find (Bloom et al. 2020, Jones 2022). The knowledge law of motion (9) becomes

$$\mathrm{d} \ln \boldsymbol{q}_t / \mathrm{d}t = \lambda \cdot (\ln \boldsymbol{\eta} + \ln \boldsymbol{s}_t + \bar{g}t + (\boldsymbol{\Omega} - (1+\kappa) \boldsymbol{I}) \ln \boldsymbol{q}_t).$$

Integrating the ODE system over time, we get

$$\ln \boldsymbol{q}_t = e^{\lambda(\boldsymbol{\Omega}-(1+\kappa)\boldsymbol{I})t} \left[\ln \boldsymbol{q}_0 + \lambda \int_0^t e^{-\lambda(\boldsymbol{\Omega}-(1+\kappa)\boldsymbol{I})u} (\ln \boldsymbol{\eta} + \ln \boldsymbol{s}_u + \bar{g}u) \, \mathrm{d}u\right]$$

For given initial levels of knowledge stock and path of worker allocation, the difference in welfare under two R&D allocations $\widetilde{\boldsymbol{b}}$ and $\boldsymbol{b}$ is

$$
\begin{aligned}
& V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\widetilde{\boldsymbol{b}}\right)-V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\boldsymbol{b}\right)\\
= \; & \psi\boldsymbol{\beta}'\int_0^\infty e^{-\rho t}\left[\ln\boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right)-\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right)\right]\,\mathrm{d}t\\
= \; & \psi\lambda\boldsymbol{\beta}'\int_0^\infty e^{-\rho t}\left[\int_0^t e^{-\lambda((1+\kappa)\boldsymbol{I}-\boldsymbol{\Omega})(t-u)}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\,\mathrm{d}u\right]\,\mathrm{d}t\\
= \; & \psi\boldsymbol{\beta}'\left((1+\kappa)\boldsymbol{I}-\boldsymbol{\Omega}\right)^{-1}\left(\int_0^\infty e^{-\rho t}\left[\boldsymbol{I}-e^{-\lambda((1+\kappa)\boldsymbol{I}-\boldsymbol{\Omega})t}\right]\,\mathrm{d}t\right)\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho}\boldsymbol{\beta}'\frac{1}{1+\kappa}\left(\boldsymbol{I}-\frac{\boldsymbol{\Omega}}{1+\kappa}\right)^{-1}\left(\boldsymbol{I}-\frac{\rho}{\rho+\lambda+\lambda\kappa}\left(\boldsymbol{I}-\frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}\right)\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi\lambda}{\rho}\boldsymbol{\beta}'\frac{1}{\rho+\lambda+\kappa\lambda}\left(\boldsymbol{I}-\frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi\lambda}{\rho}\frac{1}{\kappa\lambda+\rho}\frac{\rho+\kappa\lambda}{\rho+\lambda+\kappa\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I}-\frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)
\end{aligned}
$$

It is easy to verify that $\boldsymbol{\gamma}'\equiv\frac{\rho+\kappa\lambda}{\rho+\lambda+\kappa\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I}-\frac{\boldsymbol{\Omega}}{1+\kappa+\rho/\lambda}\right)^{-1}$ sums to one; hence we have

$$
V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\widetilde{\boldsymbol{b}}\right)-V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\boldsymbol{b}\right)=\frac{\psi}{\rho}\frac{\lambda}{\rho+\kappa\lambda}\boldsymbol{\gamma}'\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right).
$$

Clearly $\boldsymbol{\gamma}$ is the optimal allocation, and, analogous to the argument in Section A.6, the consumption-equivalent welfare impact of adopting the optimal allocation is $\mathcal{L}\left(\boldsymbol{b}\right)=\exp\left(\frac{\psi\lambda}{\rho+\kappa\lambda}\boldsymbol{\gamma}'\left(\ln\boldsymbol{\gamma}-\ln\boldsymbol{b}\right)\right)$.

B.4 An Illustrative Decentralized Equilibrium

In an innovation network, knowledge is a public good, as knowledge creation benefits subsequent R&D in other sectors and all future periods. To the extent that innovators do not fully internalize such future benefits, a decentralized market does not implement the optimal R&D allocation. To demonstrate the potential inefficiency, in this section we construct a decentralized equilibrium in which innovators conduct R&D only in pursuit of profits, disregarding any beneficial spillovers their R&D activities may provide in the future. As we show, the decentralized allocation of R&D resources follows the consumption elasticities $\boldsymbol{\beta}$ along a BGP, which can be efficient only if the society is completely myopic ($\rho/\lambda\to\infty$).

It is important to note that our decentralized equilibrium lacks many real-world features of the market for innovation (e.g., multi-sector firms, mergers and acquisitions, and patent licensing). This is intentional: the goal of this section is not to capture quantitative realism but to illustrate as clearly as possible the potential inefficiency of decentralized R&D decisions given knowledge spillovers. By comparing the R&D allocations in the data to the first-best, our notion of allocative efficiency—measured by the consumption-equivalent welfare impact of reallocating R&D optimally—does not require that we take a stance on firms' equilibrium conduct; instead, it has the advantage of directly calculating the welfare impact of reallocating R&D based on the economic environment.

We demonstrate the inefficiency of decentralized equilibrium in the economic environment with general functional forms as described in Section 2.5. Specifically, consumption aggregator is $y_t = \mathcal{Y}(\{y_{it}\})$ constant-returns-to-scale, and corss-sector knowledge spillover follows a general function $\mathcal{X}_i(\{q_{jt}\})$ satisfying homogeneous-of-degree-zero with spillovers positive across sectors ($\partial \ln \mathcal{X}_i(\cdot) / \partial \ln q_{jt} > 0$ for $i \neq j$) and bounded above ($|\partial \ln \mathcal{X}_i(\cdot) / \partial \ln q_{jt}| \leq 1 \forall i, j$). We show below in Proposition 8 that along a decentralized BGP, R&D allocation follows $\ell_i/\bar{\ell} = \beta_i \equiv \frac{\partial \ln \mathcal{Y}(\{y_{it}\})}{\partial \ln y_{it}}$, which is generically inefficient given Proposition 5.

Specifically, suppose each sectoral good consists of a continuum of intermediate varieties:

$$\ln y_{it} = \int_0^1 \ln \left[q_{it}^{\psi}(\nu)\, y_{it}(\nu) \right] \mathrm{d}\nu \tag{A11}$$

where each intermediate variety is produced by a distinct monopolist one-for-one from labor. Different vintages of the same variety are perfect substitutes. Because the most recent vintage's quality is e^{λ} proportionally higher than the next best vintage, the monopolist conducts limit pricing and charges a markup $e^{\lambda\psi}$.[11] No vintages with dominated quality are produced in equilibrium.

In each sector, innovation is carried out by a large research intermediary ("R&D firm"), who hire scientists to conduct R&D and generate new innovations with Poisson arrival rate $\phi(s_{it}\mathcal{X}_i(\{q_{jt}\}))$. Upon a successful innovation, the R&D firm patents the innovation and sells the patent to a producer, who pays for the full value of the patent and becomes the monopolist of that variety until being replaced by another monopolist when a future successful innovation occurs. The law of motion for a sector's knowledge stock is

$$\frac{\mathrm{d} \ln q_{it}}{\mathrm{d}t} = \widetilde{f}(s_{it}\mathcal{X}_i(\{q_{jt}\})), \quad \text{where } \widetilde{f}(\cdot) \equiv \lambda \times \phi(\cdot).$$

The representative consumer receives all workers' and scientists' income and profits of producers and the R&D firms. Given the initial state variables $\{q_{i0}\}_{i=1}^{K}$, a decentralized equilibrium is the time path of prices, quantities, and knowledge stocks such that production firms set prices to maximize profits, the consumer chooses bundles of goods to consume to maximize utility, and potential entrants hire scientists for R&D to maximize expected profits. A decentralized BGP is an equilibrium in which all sectors' knowledge stock grows at the same constant rate.

Note that when worker allocation is constant across varieties in each sector, $\ell_{it}(\nu) = \ell_{it} \quad \forall \nu$—which is true in the decentralized equilibrium, as shown in Proposition 8 below—the economic environment described here coincides with that in Section 2.5. Following Section 2.5, we let $\beta_i \equiv \frac{\partial \ln \mathcal{Y}(\{y_{it}\})}{\partial \ln y_{it}}$ denote the consumption elasticity with respect to sectoral good i along a decentralized BGP. This is also the consumer expenditure share on good i.

Proposition 8. *In the decentralized BGP, the allocations of R&D and production resources both follow the consumption elasticities:* $\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell}$ *and* $s_{it} = \beta_i \bar{s}$.

Proof. We normalize the consumer price index to one for all times t. The consumer spends fraction

[11] Note that λ is proportional to the profit margin.

β_i of their income on sectoral composite good i, with

$$p_{it} y_{it} = \beta_i y_t \quad \text{for all } i, t. \tag{A12}$$

The sectoral composite aggregator (A11) further implies that the total revenue of each variety ν is also equal to $\beta_i y_t$, and, because each monopolist sets a markup $e^{\lambda \psi}$, we derive the profits in each sector i as

$$\pi_{it}(\nu) = \left(1 - e^{-\lambda \psi}\right) \beta_i y_t \quad \text{for all } i, t, \nu. \tag{A13}$$

Because all varieties have identical markups, the worker allocation is identical across varieties within each sector. Given a constant markup across all sectors, the total worker allocation in each sector is also proportional to the consumption shares β_i:

$$\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell} \quad \text{for all } i, t, \nu. \tag{A14}$$

Along the BGP, a monopolist in each sector has the same Poisson rate $\bar{\phi}$ to be replaced by an innovating entrant. The value of a monopolistic firm is thus

$$v_{it} \equiv \int_t^\infty e^{-\left(r + \bar{\phi}\right)(s-t)} \pi_{is} \, \mathrm{d}s, \tag{A15}$$

where r is the interest rate. Note we have suppressed the index for variety since all varieties have the same profits and thus the same value within each sector. Because sectoral profits are always proportional to the consumption shares at all times, we have

$$v_{it} / v_{jt} = \beta_i / \beta_j \quad \text{for all } i, j, t. \tag{A16}$$

Entrants hire scientists to conduct research in order to become future monopolists. The marginal value from an additional scientist must be equalized across sectors, further implying

$$v_{it} \frac{\partial \phi\left(s_{it} \mathcal{X}_i\left(\{q_{kt}\}\right)\right)}{\partial s_{it}} = w_t^s$$

where w_t^s is the wage rate of a scientist at time t. Using the fact that $v_{it}/v_{jt} = \beta_i/\beta_j$, we have

$$\frac{\beta_i}{s_{it}} \frac{\partial \phi\left(s_{it} \mathcal{X}_i\left(\{q_{kt}\}\right)\right)}{\partial \ln s_{it}} = \frac{\beta_j}{s_{jt}} \frac{\partial \phi\left(s_{jt} \mathcal{X}_j\left(\{q_{kt}\}\right)\right)}{\partial \ln s_{jt}}$$

Further note that $\frac{\partial \phi(s_{jt} \mathcal{X}_j(\{q_{kt}\}))}{\partial \ln s_{jt}} = \phi' \cdot \mathcal{X}_j\left(\{q_{kt}\}\right) s_{jt}$ which must be the same across all sectors along a BGP. Hence we obtain that scientist allocation must also follow the consumption share, that is, $s_{it}/\bar{s} = \beta_i$ for all t, as desired. □

Intuitively, varieties in a sector with higher consumption share β_i have proportionally higher revenue, employment, and flow profits. Since the rate at which an innovating entrant replaces a producing monopolist is the same across all sectors along a BGP, a monopolistic firm's value is also proportional to the consumption share β_i of the sector. Because entrants conduct research to obtain that monopolistic value, the marginal value from an additional scientist must be equalized across sectors, and the innovation production function (4) thus implies that R&D allocation must follow $s_{it} = \beta_i \bar{s}$ along the BGP.

According to Proposition 5, a necessary condition for the decentralized balanced growth path to be efficient is for the society to be myopic (with $\rho \to \infty$). While the social planner takes into account both R&D's direct effect on product quality as well as the infinite rounds of indirect network spillover effects, the decentralized allocation is driven by firm profits and thus accounts only for the direct effect, as infinitesimal firms cannot monetize the future spillover effects of their own R&D.

B.5 Constrained Optimal R&D Allocations

In some settings, for instance under political or feasibility constraints, a planner may only be able to reallocate resources across a subset $\mathcal{K} \subset \{1, \ldots, K\}$ of sectors. We now generalize our results to such an environment. We show that our earlier results extend naturally: resources among sectors in $\mathcal{K}$ should be allocated proportionally to the unconstrained optimal allocation $\boldsymbol{\gamma}$. We generalize the welfare sufficient statistic to this setting as well.

For a generic allocation vector $\boldsymbol{b}$, we denote $\boldsymbol{b}^{\mathcal{K}}$ as the $|\mathcal{K}| \times 1$ allocation vector that sums to one with entries proportional to $\boldsymbol{b}$ for all sectors in $\mathcal{K}$ (i.e., $b_i^{\mathcal{K}} \equiv \frac{b_i}{\sum_{j \in \mathcal{K}} b_j}$ for $i \in \mathcal{K}$).

Proposition 9. *Suppose R&D allocations in sectors $k \notin \mathcal{K}$ are given exogenously and that the planner can only choose R&D allocations in sectors $k \in \mathcal{K}$ when solving the planning problem in (7). Along the entire equilibrium path, the constrained optimal R&D allocation is $s_i = \gamma_i^{\mathcal{K}} \left(\bar{s} - \sum_{k \notin \mathcal{K}} s_k\right)$ for $i \in \mathcal{K}$. The consumption-equivalent welfare gains from adopting the constrained-optimal R&D allocation (instead of allocation $\boldsymbol{b}$) is $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\psi\lambda}{\rho}\left(\sum_{j \in \mathcal{K}} \gamma_j\right)\left(\boldsymbol{\gamma}^{\mathcal{K}}\right)'\left(\ln \boldsymbol{\gamma}^{\mathcal{K}} - \ln \boldsymbol{b}^{\mathcal{K}}\right)\right)$.*

The Proposition shows that among sectors in which the planner can allocate resources, the constrained-optimal resource allocation is proportional to the unconstrained-optimal allocation $\boldsymbol{\gamma}$. For the welfare sufficient statistic, note that the relative entropy of $\boldsymbol{b}^{\mathcal{K}}$ from $\boldsymbol{\gamma}^{\mathcal{K}}$, $\left(\boldsymbol{\gamma}^{\mathcal{K}}\right)'\left(\ln \boldsymbol{\gamma}^{\mathcal{K}} - \ln \boldsymbol{b}^{\mathcal{K}}\right)$, summarizes the distance relative to the first-best allocation among sectors in $\mathcal{K}$. Relative to the welfare formula (15) for the unconstrained optimal allocation, the new term $\sum_{j \in \mathcal{K}} \gamma_j \leq 1$ (with equality when $\mathcal{K}$ includes all sectors) reflects the fact that there is less to be gained when the planner can reallocate resources across fewer sectors.

Proof. Let $s^{\mathcal{K}} \equiv \bar{s} - \sum_{k \notin \mathcal{K}} s_k$ denote the available resource the planner can allocate among sectors in $\mathcal{K}$, and let $\gamma_i^{\mathcal{K}}$ denote the constrained-optimal share of $s^{\mathcal{K}}$ allocated to sector i. That $\gamma_i^{\mathcal{K}}$ is time-invariant follows from the same proof as Proposition 3. $\boldsymbol{\gamma}^{\mathcal{K}}$ is thus the solution to

$$\boldsymbol{\gamma}^{\mathcal{K}} = \arg \max_{\{\delta_i\}_{i \in \mathcal{K}}} \sum_{i \in \mathcal{K}} \gamma_i \left(\ln \delta_i - \ln b_i\right) \quad \text{s.t.} \sum_{i \in \mathcal{K}} \delta_i = 1.$$

It is thus immediate that $\gamma_i^{\mathcal{K}} = \frac{\gamma_i}{\sum_{j \in \mathcal{K}} \gamma_j}$. By Proposition 3, the welfare gains from adopting the constrained optimal allocation is

$$\frac{\psi\lambda}{\rho^2}\left(\sum_{i \in \mathcal{K}} \gamma_i \left(\ln \gamma_i^{\mathcal{K}}\left(\sum_{i \in \mathcal{K}} b_i\right) - \ln b_i^{\mathcal{K}}\left(\sum_{i \in \mathcal{K}} b_i\right)\right) + \sum_{i \notin \mathcal{K}} \gamma_i \left(\ln b_i - \ln b_i\right)\right),$$

the consumption-equivalent gains then simplifies to the formula in the Proposition. □

The Proposition also holds in an environment with foreign spillovers, which we state below.

Proposition 10. *Consider an open economy with R&D self-sufficiency ξ and given paths of foreign knowledge and relative import prices $\left\{\boldsymbol{q}_t^f, p_t^f\right\}_{t=0}^{\infty}$. Suppose R&D allocations in sectors $k \notin \mathcal{K}$ are given exogenously and that the planner can only choose R&D allocations in sectors $k \in \mathcal{K}$ when solving the planning problem in (20). Along the entire equilibrium path, the constrained optimal R&D allocation is $s_i = \gamma_i^{\mathcal{K}}\left(\bar{s} - \sum_{k \notin \mathcal{K}} s_k\right)$ for $i \in \mathcal{K}$. The consumption-equivalent welfare gains from adopting the constrained-optimal R&D allocation (instead of allocation $\boldsymbol{b}$) is $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\psi\lambda}{\rho}\xi\left(\sum_{j\in\mathcal{K}}\gamma_j\right)\left(\boldsymbol{\gamma}^{\mathcal{K}}\right)'\left(\ln\boldsymbol{\gamma}^{\mathcal{K}} - \ln\boldsymbol{b}^{\mathcal{K}}\right)\right)$.*

B.6 Optimal R&D Allocation in Large Open Economies

In the open economy environment presented in the main text, we studied the problem of a domestic planner who takes the paths of import prices and foreign knowledge as given. In this appendix section, we construct an environment in which a domestic planner internalizes the impact of domestic allocations on foreign variables. This analysis is empirically relevant for studying the R&D allocation in the U.S., a country that generates significant knowledge spillovers to other economies.

Consider an environment with two economies, home (U.S.) and foreign (rest of the world). The home consumer has preferences

$$V = \int_0^\infty e^{-\rho t}\left(\sigma^h \ln c_t^{hh} + \left(1-\sigma^h\right)\ln c_t^{hf}\right)\,\mathrm{d}t, \tag{A17}$$

where c_t^{hh} is the home consumption of home goods and c_t^{hf} is the home consumption of foreign goods. Home goods is a Cobb-Douglas aggregator over sectoral composite goods, which are produced from labor (equations 2 and 3). We can simplify the home production functions as

$$\ln y_t^h = \sum_i \beta_i\left(\psi \ln q_{it}^h + \ln \ell_{it}^h\right). \tag{A18}$$

Home can import the foreign goods c_t^{hf} by exporting unconsumed home goods $\left(y_t^h - c_t^{hh}\right)$. Home innovation production function follows

$$n_{it}^h = s_{it}^h \chi_{it}^h, \qquad \text{where } \chi_{it}^h = \eta_i^h \prod_{j=1}^K\left[\left(q_{jt}^h\right)^{x_{ij}^h}\left(q_{jt}^f\right)^{1-x_{ij}^h}\right]^{\omega_{ij}}, \tag{A19}$$

and the law of motion for home knowledge stock is

$$\frac{\mathrm{d}\ln q_{it}^h}{\mathrm{d}t} = \lambda \ln\left(n_{it}^h / q_{it}^h\right). \tag{A20}$$

Home is endowed with workers $\bar{\ell}^h$ and scientists $\bar{s}^h$. The foreign economy has analogous preferences and technologies, swapping superscripts h and f.

We study the home planner's problem of allocating workers and scientists to maximize home welfare, while taking the time path of foreign allocations $\left\{\boldsymbol{\ell}_t^f, \boldsymbol{s}_t^f\right\}$ as given and decentralizing international trade. Given home and foreign output y_t^h, y_t^f, Cobb-Douglas preferences imply that the home consumer spends $\left(1-\sigma^h\right)$ fraction of income on home imports, and that the foreign consumer spends $\left(1-\sigma^f\right)$ fraction of income on home exports. Trade balance therefore implies that home consumption of foreign goods is $\left(1-\sigma^f\right) y_t^f$. Hence, given flow output y_t^h, y_t^f, the home consumer's flow utility is

$$\sigma^h \ln c_t^{hh} + \left(1-\sigma^h\right) \ln c_t^{hf} = \sigma^h \ln \sigma^h y_t^h + \left(1-\sigma^h\right) \ln \left(1-\sigma^f\right) y_t^f.$$

Substituting into (A17), we can write the home planning problem as

$$V^*\left(\left\{\boldsymbol{\ell}_t^f, \boldsymbol{s}_t^f\right\}_{t=0}^{\infty}\right) \equiv \max_{\left\{s_{it}^h, \ell_{it}^h\right\}} \int_0^\infty e^{-\rho t}\left(\sigma^h \ln y_t^h + \left(1-\sigma^h\right) \ln y_t^f\right) \mathrm{d}t, \tag{A21}$$

subject to the innovation production functions (A20 and A19), goods production function (A18), and the corresponding foreign innovation and goods production functions

$$\frac{\mathrm{d} \ln q_{it}^f}{\mathrm{d}t} = \ln \eta_i^f + \ln s_{it}^f + \sum_{j=1}^K \omega_{ij}\left(x_{ij}^f \ln q_{jt}^f + \left(1-x_{ij}^f\right) \ln q_{jt}^h\right),$$

$$\ln y_t^f = \sum_i \beta_i \left(\ln q_{it}^f + \ln \ell_{it}^f\right),$$

with market clearing conditions $\sum_i s_{it}^h = \bar{s}^h$ and $\sum_i \ell_{it}^h = \bar{\ell}^h$.

To solve the home planner's problem, first consider a hypothetical world as an integrated economy in which resources can freely move across countries, and where the home planner can choose worker and scientist allocations in both economies; then, our closed economy analysis in Section 2.2 exactly applies: the solution would be characterized exactly by our closed economy results in Lemma 1 and Proposition 1, recognizing that there are $K \times 2$ sectors in both economies, with home's consumption elasticity captured by

$$\hat{\boldsymbol{\beta}}' \equiv \left[\sigma^h \boldsymbol{\beta}', \left(1-\sigma^h\right) \boldsymbol{\beta}'\right], \tag{A22}$$

and the innovation network captured by

$$\hat{\boldsymbol{\Omega}} \equiv \begin{bmatrix} \boldsymbol{\Omega} \circ \boldsymbol{X}^h & \boldsymbol{\Omega} - \boldsymbol{\Omega} \circ \boldsymbol{X}^h \\ \boldsymbol{\Omega} - \boldsymbol{\Omega} \circ \boldsymbol{X}^f & \boldsymbol{\Omega} \circ \boldsymbol{X}^f \end{bmatrix}. \tag{A23}$$

Optimal worker allocation should follow $\hat{\boldsymbol{\beta}}$, and optimal R&D allocation should follow

$$\hat{\gamma}' \equiv \frac{\rho}{\rho+\lambda} \hat{\boldsymbol{\beta}}' \left(\boldsymbol{I}_{2K \times 2K} - \frac{\hat{\boldsymbol{\Omega}}}{1+\rho/\lambda}\right)^{-1}. \tag{A24}$$

Next, recognize that the actual home planner's problem (A21) is essentially the same as in the hypothetical integrated economy, but with the additional constraint that the home planner can only allocate resources domestically. We can apply the result in Section B.5 to get the following Proposition.

Proposition 11. *The optimal resource allocation for an open economy planner who takes the path of foreign allocations* $\left\{\boldsymbol{\ell}_t^f, \boldsymbol{s}_t^f\right\}$ *as given and solves the problem in (A21) is to allocate workers according to* $\hat{\boldsymbol{\beta}}^{\mathcal{K}}$ *(i.e.,* $\ell_{it}^h/\bar{\ell}^h = \hat{\beta}_i^{\mathcal{K}}$*) and R&D resources according to* $\hat{\boldsymbol{\gamma}}^{\mathcal{K}}$ *(i.e.,* $s_{it}^h/\bar{s}^h = \hat{\gamma}_i^{\mathcal{K}}$*), where* $\mathcal{K}$ *is the set of domestic sectors, and*

$$\hat{\beta}_i^{\mathcal{K}} = \frac{\hat{\beta}_i}{\sum_{j\in\mathcal{K}}\hat{\beta}_j}, \qquad \hat{\gamma}_i^{\mathcal{K}} = \frac{\hat{\gamma}_i}{\sum_{j\in\mathcal{K}}\hat{\gamma}_j}.$$

The consumption-equivalent welfare gains from adopting the optimal domestic R&D allocation (instead of allocation $\boldsymbol{b}$*) is* $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\psi\lambda}{\rho}\left(\sum_{j\in\mathcal{K}}\hat{\gamma}_j\right)\left(\hat{\boldsymbol{\gamma}}^{\mathcal{K}}\right)'\left(\ln\hat{\boldsymbol{\gamma}}^{\mathcal{K}} - \ln\boldsymbol{b}\right)\right)$.

B.7 General Functional Forms and Endogenous Innovation Network with Foreign Spillovers

We now extend our analysis in Section 2.6 to incorporate general functional forms, thereby endogenizing the degree to which domestic innovation benefits from foreign spillovers. We show, analogous to our closed-economy analysis in Section 2.5, that Proposition 7 in the main text continues to hold, as a first-order approximation around a balanced growth path, to the welfare impact of adopting the optimal R&D allocation.

For completeness, we provide all equations to this economic environment:

$$V\left(\left\{q_{jt}^f, p_t^f\right\}\right) = \int_0^\infty e^{-\rho t}\ln\mathcal{C}\left(c_t^d, c_t^f\right)\,\mathrm{d}t,$$

$$p_t^f c_t^f = y_t - c_t^d.$$

$$y_t = \mathcal{Y}\left(\left\{q_{it}^{\psi}\ell_{it}\right\}\right)$$

$$\mathrm{d}\ln q_{it}/\,\mathrm{d}t = \lambda\cdot\left(\ln\left(b_{it}\bar{s}\right) + \ln\mathcal{X}_i\left(\left\{q_{jt}, q_{jt}^f\right\}\right)\right)$$

The first equation represents consumer welfare; the second equation is trade balance; the third equation is the production function; the last equation is the law of motion for sectoral knowledge stock. The function $\mathcal{X}_i\left(\left\{q_{jt}, q_{jt}^f\right\}\right)$ captures how domestic innovation in sector i benefits from domestic and foreign knowledge; it is a generalization of the Cobb-Douglas functional form in equation (19). We assume $\mathcal{C}$ and $\mathcal{Y}$ are constant-returns-to-scale, and that $\mathcal{X}_i(\cdot)$ is homogeneous-of-degree-zero, $\partial\ln\mathcal{X}_i(\cdot)/\partial\ln q_{jt} \geq 0\,\forall i\neq j$, $\partial\ln\mathcal{X}_i(\cdot)/\partial\ln q_{jt}^f\,\forall i,j$, and $\left|\partial\ln\mathcal{X}_i(\cdot)/\partial\ln q_{jt}\right| \leq 1\forall i,j$.

Consider the economy initially at $t=0$ in a BGP with R&D allocation $\boldsymbol{b}$, where foreign knowledge q_{jt}^f grows at exogenous rate g in all sectors, and p_t^f is time-invariant. Define

$$\beta_i \equiv \left.\frac{\partial\ln\mathcal{Y}\left(\{y_{it}\}\right)}{\partial\ln y_{it}}\right|_{t=0}, \qquad \theta_{ij} \equiv \begin{cases} \left.\frac{\partial\ln\mathcal{X}_i\left(\left\{q_{it}, q_{jt}^f\right\}\right)}{\partial\ln q_{jt}}\right|_{t=0} & \text{if } i=j \\ 1 + \left.\frac{\partial\ln\mathcal{X}_i\left(\left\{q_{it}, q_{jt}^f\right\}\right)}{\partial\ln q_{jt}}\right|_{t=0} & \text{otherwise.} \end{cases}$$

$\boldsymbol{\beta} \equiv [\beta_i]$ and $\boldsymbol{\Theta} \equiv [\theta_{ij}]$ are the consumption and innovation spillover elasticities with respect to domestic knowledge stock evaluated in the initial BGP. Note that $\boldsymbol{I}-\boldsymbol{\Theta}$ is generically invertible, as the economy features aggregate decreasing-returns-to-scale with respect to domestic knowledge stock.

The Gateaux derivative of welfare with respect to log R&D allocation in the direction of $\boldsymbol{h}$ is $\frac{\lambda}{\rho}\xi\boldsymbol{\gamma}'\boldsymbol{h}$; the proof parallels that of Proposition 5.

B.8 Sector-Specific λ_i's

We now introduce a theoretical extension allowing for sector-specific λ_i. Let $\boldsymbol{\Lambda} \equiv \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_K \end{bmatrix}$ denote the diagonal matrix with λ_i along the diagonal, and let $\boldsymbol{\lambda} \equiv [\lambda_i]$ denote the vector of λ_i's. We show the optimal R&D allocation $\boldsymbol{\gamma}$ should follow (scaled so that $\boldsymbol{\gamma}$ sums to one)

$$\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \boldsymbol{\Omega} + \rho\boldsymbol{\Lambda}^{-1}\right)^{-1}$$

and the consumption-equivalent welfare impact of adopting the optimal allocation is

$$\mathcal{L}\left(\boldsymbol{b}\right) = \exp\left(\psi\boldsymbol{\beta}'\left(\boldsymbol{I} - \boldsymbol{\Omega} + \rho\boldsymbol{\Lambda}^{-1}\right)^{-1}\left(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}\right)\right).$$

Specifically, the social planner's problem is

$$\max_{\{\gamma_t\}\text{ s.t. }\gamma_t'1=1\forall t} \int_0^\infty e^{-\rho t}\boldsymbol{\beta}' \ln\boldsymbol{q}_t \, \mathrm{d}t$$
$$\text{s.t. } \mathrm{d}\ln\boldsymbol{q}_t/\mathrm{d}t = \boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{\gamma}_t + (\boldsymbol{\Omega} - \boldsymbol{I})\ln\boldsymbol{q}_t\right) \tag{A25}$$

The control variable is $\boldsymbol{\gamma}_t$ and the state variable is $\boldsymbol{q}_t$. Denote the co-state variables as $\boldsymbol{\mu}_t$. The current-value Hamiltonian is

$$H(\gamma_t, \boldsymbol{q}_t, \boldsymbol{\mu}_t, \zeta) = \boldsymbol{\beta}'\ln\boldsymbol{q}_t + \boldsymbol{\mu}_t'\boldsymbol{\Lambda}\left[\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{\gamma}_t + (\boldsymbol{\Omega} - \boldsymbol{I})\ln\boldsymbol{q}_t\right] + \zeta\left(1 - \boldsymbol{\gamma}_t'\boldsymbol{1}\right).$$

For notational simplicity we suppress dependence on time for control, state, and co-state variables:

$$H(\{\gamma_i\}, \{q_i\}, \{\mu_i\}, \zeta, t) = \sum_i \beta_i \ln q_i + \sum_i \mu_i\lambda_i\left(\ln\eta_i + \ln\bar{s} + \ln\gamma_i + \sum_j \omega_{ij}\ln q_j - \ln q_i\right) + \zeta(1 - \sum_i \gamma_i).$$

By the maximum principle

$$H_{\gamma_i} = 0 \Longleftrightarrow \frac{\lambda_i\mu_i}{\gamma_i} = \zeta \quad \forall i \tag{A26}$$

$$H_{\ln q_i} = \rho\mu_i - \dot{\mu}_i \Longleftrightarrow \beta_i - \lambda_i\mu_i + \sum_j \lambda_j\mu_j\omega_{ji} = \rho\mu_i - \dot{\mu}_i \tag{A27}$$

Similar to the proof of Proposition 1, we can show $\dot{\mu}_i = 0$ for all i; hence,

$$\boldsymbol{\gamma}' \propto \boldsymbol{\mu}'\boldsymbol{\Lambda},$$

$$\begin{aligned} \boldsymbol{\beta}' &= \boldsymbol{\mu}'\left((\rho + \boldsymbol{\Lambda})\boldsymbol{I} - \boldsymbol{\Lambda}\boldsymbol{\Omega}\right) \\ &= \boldsymbol{\mu}'\boldsymbol{\Lambda}\left(\rho\boldsymbol{\Lambda}^{-1} + \boldsymbol{I} - \boldsymbol{\Omega}\right) \\ &\propto \boldsymbol{\gamma}'\left(\rho\boldsymbol{\Lambda}^{-1} + \boldsymbol{I} - \boldsymbol{\Omega}\right) \end{aligned}$$

Hence

$$\boldsymbol{\gamma}' = \boldsymbol{\beta}'\left(\boldsymbol{I} - \boldsymbol{\Omega} + \rho\boldsymbol{\Lambda}^{-1}\right)^{-1}.$$

To derive the welfare impact of R&D reallocation, let $g_i^q \equiv \frac{\mathrm{d}\ln q_{it}}{\mathrm{d}t}$ be the growth rate of knowledge stock in sector i along the BGP. We know

$$\boldsymbol{g}^q = \boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I})\ln\boldsymbol{q}_t\right)$$

Take derivative with respect to time,

$$\boldsymbol{0} = \boldsymbol{\Lambda}\left(\boldsymbol{\Omega} - \boldsymbol{I}\right)\frac{\mathrm{d}\ln\boldsymbol{q}_t}{\mathrm{d}t}$$

So that

$$\boldsymbol{g}^q = \boldsymbol{\Omega}\boldsymbol{g}^q$$

We know the only right-Perron eigenvector of $\boldsymbol{\Omega}$ is the constant vector; hence all sectors must grow at the same rate g^q, satisfying

$$\begin{aligned} g^q\boldsymbol{1} &= \boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I})\ln\boldsymbol{q}_t\right) \\ \implies g^q\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\boldsymbol{1} &= \boldsymbol{a}'\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I})\ln\boldsymbol{q}_t\right) \\ \implies g^q &= \frac{\boldsymbol{a}'\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{b}\right)}{\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\boldsymbol{1}} \end{aligned}$$

Let $\boldsymbol{A} \equiv \frac{\boldsymbol{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\boldsymbol{1}}$. Note $(\boldsymbol{I} - \boldsymbol{A})\ln\bar{s}\boldsymbol{1} = \boldsymbol{0}$, and that

$$\begin{aligned} (\boldsymbol{\Omega} - \boldsymbol{I}) &= (\boldsymbol{\Omega} - \boldsymbol{I})(\boldsymbol{I} - \boldsymbol{A}) \\ &= -(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A})(\boldsymbol{I} - \boldsymbol{A}) \end{aligned}$$

Let $\widetilde{\ln\boldsymbol{q}_t} \equiv (\boldsymbol{I} - \boldsymbol{A})\ln\boldsymbol{q}_t$; then

$$\begin{aligned} (\boldsymbol{I} - \boldsymbol{A})\,\mathrm{d}\ln\boldsymbol{q}_t/\mathrm{d}t &= \left(\boldsymbol{\Lambda} - \frac{\boldsymbol{1}\boldsymbol{a}'}{\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\boldsymbol{1}}\right)\left(\ln\boldsymbol{\eta} + \ln\bar{s}\boldsymbol{1} + \ln\boldsymbol{b}\right) + \boldsymbol{\Lambda}\left(\boldsymbol{\Omega} - \boldsymbol{I}\right)\ln\boldsymbol{q}_t \\ &= (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta} + \ln\boldsymbol{b}\right) - \boldsymbol{\Lambda}\left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)(\boldsymbol{I} - \boldsymbol{A})\ln\boldsymbol{q}_t \end{aligned}$$

$$\frac{\mathrm{d}\widetilde{\ln\boldsymbol{q}_t}}{\mathrm{d}t} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta} + \ln\boldsymbol{b}\right) - \boldsymbol{\Lambda}\left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}\right)(\boldsymbol{I} - \boldsymbol{A})\widetilde{\ln\boldsymbol{q}_t}$$

We can integrate the ODE system:

$$\begin{aligned}
\widetilde{\ln \boldsymbol{q}_t} &= e^{-\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\left[\widetilde{\ln \boldsymbol{q}_0}+\int_0^t e^{\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})s}\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)\,\mathrm{d}s\right] \\
&= e^{-\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\widetilde{\ln \boldsymbol{q}_0}+\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left[\boldsymbol{I}-e^{-\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right]\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta}+\ln\boldsymbol{b}\right)
\end{aligned}$$

We know

$$\begin{aligned}
\boldsymbol{A}\frac{\mathrm{d}\ln\boldsymbol{q}_t}{\mathrm{d}t} &= \frac{\mathbf{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\mathbf{1}}\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta}+\ln\bar{s}\mathbf{1}+\ln\boldsymbol{b}+\left(\boldsymbol{\Omega}-\boldsymbol{I}\right)\ln\boldsymbol{q}_t\right) \\
&= \frac{\mathbf{1}\boldsymbol{a}'}{\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\mathbf{1}}\left(\ln\boldsymbol{\eta}+\ln\bar{s}\mathbf{1}+\ln\boldsymbol{b}\right) \\
&= \boldsymbol{A}\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta}+\ln\bar{s}\mathbf{1}+\ln\boldsymbol{b}\right)
\end{aligned}$$

Hence

$$\boldsymbol{A}\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right)=\boldsymbol{A}\ln\boldsymbol{q}_0+\boldsymbol{A}\boldsymbol{\Lambda}\left(\ln\boldsymbol{\eta}+\ln\bar{s}\mathbf{1}+\ln\boldsymbol{b}\right)t$$

Now consider starting from the same initial knowledge stock $\boldsymbol{q}_0$ but with two different time-invariant R&D allocations $\widetilde{\boldsymbol{b}}$ and $\boldsymbol{b}$,

$$\boldsymbol{A}\ln\boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right)-\boldsymbol{A}\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right)=\boldsymbol{A}\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)t$$

we have the following difference in knowledge stock over time:

$$\begin{aligned}
\ln\boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right)-\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right) &= \boldsymbol{A}\ln\boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right)-\boldsymbol{A}\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right) \\
&\quad+\widetilde{\ln\boldsymbol{q}_t}\left(\widetilde{\boldsymbol{b}}\right)-\widetilde{\ln\boldsymbol{q}_t}\left(\boldsymbol{b}\right) \\
&= \left[\boldsymbol{A}\boldsymbol{\Lambda}t+\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left[\boldsymbol{I}-e^{-\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right]\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\right]\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)
\end{aligned}$$

The difference in consumer welfare under two time-invariant paths of R&D allocations is

$$\begin{aligned}
& V\left(\boldsymbol{q}_0;\left\{\boldsymbol{\ell}_t\right\},\widetilde{\boldsymbol{b}}\right)-V\left(\boldsymbol{q}_0;\left\{\boldsymbol{\ell}_t\right\},\boldsymbol{b}\right)\\
= \; & \psi\boldsymbol{\beta}'\int_0^\infty e^{-\rho t}\left[\ln\boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right)-\ln\boldsymbol{q}_t\left(\boldsymbol{b}\right)\right]\,\mathrm{d}t\\
= \; & \psi\boldsymbol{\beta}'\int_0^\infty e^{-\rho t}\left[\boldsymbol{A}\boldsymbol{\Lambda}t+\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left[\boldsymbol{I}-e^{-\boldsymbol{\Lambda}(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A})t}\right]\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\right]\,\mathrm{d}t\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho^2}\boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
& +\psi\boldsymbol{\beta}'\boldsymbol{\Lambda}^{-1}\left(\boldsymbol{I}-\boldsymbol{\Omega}+\boldsymbol{A}\right)^{-1}\left[\frac{1}{\rho}\boldsymbol{I}-\int_0^\infty\left(e^{-((\rho\boldsymbol{I}+\boldsymbol{\Lambda})\boldsymbol{I}-\boldsymbol{\Lambda}(\boldsymbol{\Omega}-\boldsymbol{A}))t}\right)\,\mathrm{d}t\right]\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho}\boldsymbol{\beta}'\left\{\frac{1}{\rho}\boldsymbol{A}\boldsymbol{\Lambda}+\left(\left[\left(\rho\boldsymbol{I}+\boldsymbol{\Lambda}\right)-\boldsymbol{\Lambda}\left(\boldsymbol{\Omega}-\boldsymbol{A}\right)\right]\right)^{-1}\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\right\}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho}\boldsymbol{\beta}'\left(\left[\left(\rho\boldsymbol{I}+\boldsymbol{\Lambda}\right)-\boldsymbol{\Lambda}\left(\boldsymbol{\Omega}-\boldsymbol{A}\right)\right]\right)^{-1}\left(\left[\boldsymbol{I}+\frac{1}{\rho}\boldsymbol{\Lambda}\left(\boldsymbol{I}-\left(\boldsymbol{\Omega}-\boldsymbol{A}\right)\right)\right]\right)\boldsymbol{A}\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
& +\frac{\psi}{\rho}\boldsymbol{\beta}'\left(\left[\left(\rho\boldsymbol{I}+\boldsymbol{\Lambda}\right)-\boldsymbol{\Lambda}\left(\boldsymbol{\Omega}-\boldsymbol{A}\right)\right]\right)^{-1}\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho^2}\boldsymbol{\beta}'\left(\left[\left(\rho\boldsymbol{I}+\boldsymbol{\Lambda}\right)-\boldsymbol{\Lambda}\left(\boldsymbol{\Omega}-\boldsymbol{A}\right)\right]\right)^{-1}\left(\rho\boldsymbol{I}+\boldsymbol{\Lambda}\boldsymbol{A}\right)\boldsymbol{\Lambda}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho^2}\boldsymbol{\beta}'\left(\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\boldsymbol{A}\right]\right)^{-1}\left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}\right)\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
= \; & \frac{\psi}{\rho^2}\boldsymbol{\beta}'\left(\left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}\right)^{-1}\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\boldsymbol{A}\right]\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)
\end{aligned}$$

Let $\alpha\equiv\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}\boldsymbol{1}$. Note

$$\left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}\right)^{-1}=\frac{1}{\rho}\left(\boldsymbol{I}-\frac{\alpha}{1+\alpha\rho}\boldsymbol{A}\boldsymbol{\Lambda}\right)$$

To see this,

$$\begin{aligned}
& \left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}\right)\frac{1}{\rho}\left(\boldsymbol{I}-\frac{\alpha}{1+\alpha\rho}\boldsymbol{A}\boldsymbol{\Lambda}\right)\\
= \; & \frac{1}{\rho}\left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}-\left(\rho\boldsymbol{I}+\boldsymbol{A}\boldsymbol{\Lambda}\right)\frac{\alpha}{1+\alpha\rho}\boldsymbol{A}\boldsymbol{\Lambda}\right)\\
= \; & \boldsymbol{I}+\frac{1}{\rho}\left(\frac{\boldsymbol{1}\boldsymbol{a}'}{\alpha}-\left(\rho\boldsymbol{I}+\frac{\boldsymbol{1}\boldsymbol{a}'}{\alpha}\right)\frac{1}{1+\alpha\rho}\boldsymbol{1}\boldsymbol{a}'\right)\\
= \; & \boldsymbol{I}+\frac{1}{\rho}\left(\frac{1}{\alpha}-\frac{1+\alpha\rho}{\alpha}\frac{1}{1+\alpha\rho}\right)\boldsymbol{1}\boldsymbol{a}'\\
= \; & \boldsymbol{I}
\end{aligned}$$

Hence,

$$
\begin{aligned}
& V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\widetilde{\boldsymbol{b}}\right)-V\left(\boldsymbol{q}_0;\{\boldsymbol{\ell}_t\},\boldsymbol{b}\right)\\
&=\frac{\psi}{\rho^2}\boldsymbol{\beta}'\left(\frac{1}{\rho}\left(\boldsymbol{I}-\frac{\alpha}{1+\alpha\rho}\boldsymbol{A}\boldsymbol{\Lambda}\right)\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\boldsymbol{A}\right]\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
&=\frac{\psi}{\rho}\boldsymbol{\beta}'\left(\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\frac{\boldsymbol{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\alpha}\right]-\frac{\boldsymbol{1}\boldsymbol{a}'}{1+\alpha\rho}\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\frac{\boldsymbol{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\alpha}\right]\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
&=\frac{\psi}{\rho}\boldsymbol{\beta}'\left(\left[\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}+\frac{\boldsymbol{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\alpha}\right]-\frac{\boldsymbol{1}\boldsymbol{a}'\boldsymbol{\Lambda}^{-1}}{\alpha}\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)\\
&=\frac{\psi}{\rho}\boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}\right)^{-1}\left(\ln\widetilde{\boldsymbol{b}}-\ln\boldsymbol{b}\right)
\end{aligned}
$$

Following the proof of Proposition 4, the consumption-equivalent welfare impact of adopting the optimal allocation is thus

$$\mathcal{L}(\boldsymbol{b})=\psi\boldsymbol{\beta}'\left(\boldsymbol{I}-\boldsymbol{\Omega}+\rho\boldsymbol{\Lambda}^{-1}\right)^{-1}\left(\ln\boldsymbol{\gamma}-\ln\boldsymbol{b}\right).$$

B.9 Innovation Network with Heterogeneous Row-Sums

Our baseline specification of $\boldsymbol{\Omega}$ assumes that each row sums to one (i.e., $\boldsymbol{\Omega}\boldsymbol{1}=\boldsymbol{1}$, so that $\boldsymbol{\Omega}$ is a row-stochastic Markov matrix). Because the spectral radius of any Markov matrix is equal to one, our baseline model is one with endogenous growth. The specification also motivates our measurement of the innovation network based on patent citations, $\omega_{ij}\equiv\frac{Cites_{ij}}{\sum_{k=1}^{K}Cites_{ik}}$.

In general, the knowledge spillover network is inherently difficult to measure. A reason alternative specification is to construct the network as $\omega_{ij}\propto Cites_{ij}$. This specification results in an innovation network matrix $\boldsymbol{\Omega}$ with heterogeneous row-sums ($\sum_j\omega_{ij}$ varies with i). The proportionality constant maps monotonically into the spectral radius of $\boldsymbol{\Omega}$. The model features endogenous (semi-endogenous) growth if the spectral radius is equal to (less than) one.[12]

Propositions 1 extends directly to the case where the spectral radius of $\boldsymbol{\Omega}$ is ≤ 1, as the proof does not make use of the fact that $\boldsymbol{\Omega}$ is row-stochastic. We now show Proposition 4 holds in the endogenous growth case, with the spectral radius of $\boldsymbol{\Omega}$ equal to one. Analogous results can be derived (but omitted here) in the semi-endogenous growth case as well.

Let $\boldsymbol{v}$ denote the right-Perron eigenvector of $\boldsymbol{\Omega}$, scaled so that $\boldsymbol{a}'\boldsymbol{v}=1$. Let $\boldsymbol{A}\equiv\boldsymbol{v}\boldsymbol{a}'$. We adapt the derivations in the proof of Proposition 3 to this setting, replacing $\boldsymbol{A}\equiv\boldsymbol{1}\boldsymbol{a}'$ in the baseline proof to $\boldsymbol{A}\equiv\boldsymbol{v}\boldsymbol{a}'$. Note that in the baseline setting where $\boldsymbol{\Omega}$ is row-stochastic, $\boldsymbol{v}=\boldsymbol{1}$, so the derivation below is a strict generalization.

For time-invariant R&D allocation $\boldsymbol{b}$, the law of motion of sectoral knowledge stock implies

$$\boldsymbol{a}'\frac{\mathrm{d}\ln\boldsymbol{q}}{\mathrm{d}t}=\lambda\boldsymbol{a}'\left(\ln\boldsymbol{\eta}+\ln\bar{s}\cdot\boldsymbol{1}+\ln\boldsymbol{b}\right)$$

Hence, $\boldsymbol{a}'\ln\boldsymbol{q}_t$ always grows at a constant rate (and it equals to the rate of growth along a BGP)

[12] The model features explosive growth if the spectral radius of $\boldsymbol{\Omega}$ is greater than one.

and can be solved in closed-form:

$$\boldsymbol{a}' \ln \boldsymbol{q}_t = \boldsymbol{a}' \ln \boldsymbol{q}_0 + \lambda \boldsymbol{a}' \left(\ln \boldsymbol{\eta} + \ln \bar{s} \cdot \mathbf{1} + \ln \boldsymbol{b} \right) t$$

Note that $\boldsymbol{\Omega A} = \boldsymbol{A \Omega} = \boldsymbol{A}$ and $\boldsymbol{AA} = \boldsymbol{A}$. Hence

$$\left(\boldsymbol{I} - \boldsymbol{A} \right) \left(\boldsymbol{\Omega} - \boldsymbol{I} \right) = \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A} \right) \left(\boldsymbol{I} - \boldsymbol{A} \right)$$

Left-multiply the law of motion by $(\boldsymbol{I} - \boldsymbol{A})$, substitute the above, and let $\widetilde{\ln \boldsymbol{q}_t} \equiv (\boldsymbol{I} - \boldsymbol{A}) \ln \boldsymbol{q}_t$, we get

$$\frac{\mathrm{d} \widetilde{\ln \boldsymbol{q}_t}}{\mathrm{d}t} = \lambda \left(\boldsymbol{I} - \boldsymbol{A} \right) \left(\ln \boldsymbol{\eta} + \ln \boldsymbol{b} \right) - \lambda \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A} \right) \widetilde{\ln \boldsymbol{q}_t}$$

Following the proof of Proposition 3,

$$\begin{aligned} \widetilde{\ln \boldsymbol{q}_t} &= e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \left[\widetilde{\ln \boldsymbol{q}_0} + \lambda \int_0^t e^{\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) s} \left(\boldsymbol{I} - \boldsymbol{A} \right) \left(\ln \boldsymbol{\eta} + \ln \boldsymbol{b} \right) \, \mathrm{d}s \right] \\ &= e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \widetilde{\ln \boldsymbol{q}_0} + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A} \right)^{-1} \left(\boldsymbol{I} - e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \right) \, \mathrm{d}s \left(\boldsymbol{I} - \boldsymbol{A} \right) \left(\ln \boldsymbol{\eta} + \ln \boldsymbol{b} \right) \end{aligned}$$

$$\begin{aligned} \ln \boldsymbol{q}_t &= \widetilde{\ln \boldsymbol{q}_t} + \boldsymbol{A} \ln \boldsymbol{q}_t \\ &= \boldsymbol{A} \ln \boldsymbol{q}_0 + \lambda \boldsymbol{A} \left(\ln \boldsymbol{\eta} + \ln \bar{s} \cdot \mathbf{1} + \ln \boldsymbol{b} \right) t \\ &\quad + e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \widetilde{\ln \boldsymbol{q}_0} + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A} \right)^{-1} \left(\boldsymbol{I} - e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \right) \, \mathrm{d}s \left(\boldsymbol{I} - \boldsymbol{A} \right) \left(\ln \boldsymbol{\eta} + \ln \boldsymbol{b} \right) \end{aligned}$$

Starting from the same initial knowledge stock $\boldsymbol{q}_0$ but with two different time-invariant R&D allocations $\widetilde{\boldsymbol{b}}$ and $\boldsymbol{b}$, we have the following difference in knowledge stock over time:

$$\begin{aligned} \ln \boldsymbol{q}_t \left(\widetilde{\boldsymbol{b}} \right) - \ln \boldsymbol{q}_t \left(\boldsymbol{b} \right) &= \boldsymbol{A} \ln \boldsymbol{q}_t \left(\widetilde{\boldsymbol{b}} \right) - \boldsymbol{A} \ln \boldsymbol{q}_t \left(\boldsymbol{b} \right) \\ &\quad + \widetilde{\ln \boldsymbol{q}_t} \left(\widetilde{\boldsymbol{b}} \right) - \widetilde{\ln \boldsymbol{q}_t} \left(\boldsymbol{b} \right) \\ &= \left[\boldsymbol{A} \lambda t + \left(\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A} \right)^{-1} \left(\boldsymbol{I} - e^{-\lambda (\boldsymbol{I} - \boldsymbol{\Omega} + \boldsymbol{A}) t} \right) \left(\boldsymbol{I} - \boldsymbol{A} \right) \right] \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \end{aligned}$$

The difference in consumer welfare under two time-invariant paths of R&D allocations is (derivation follows from the proof of Proposition 3)

$$\begin{aligned} & V \left(\boldsymbol{q}_0; \{ \boldsymbol{\ell}_t \}, \widetilde{\boldsymbol{b}} \right) - V \left(\boldsymbol{q}_0; \{ \boldsymbol{\ell}_t \}, \boldsymbol{b} \right) \\ =\ & \psi \boldsymbol{\beta}' \int_0^\infty e^{-\rho t} \left[\ln \boldsymbol{q}_t \left(\widetilde{\boldsymbol{b}} \right) - \ln \boldsymbol{q}_t \left(\boldsymbol{b} \right) \right] \, \mathrm{d}t \\ =\ & \frac{\psi}{\rho} \frac{\lambda}{\rho + \lambda} \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\lambda}{\rho + \lambda} \boldsymbol{\Omega} \right)^{-1} \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\ =\ & \frac{\psi \lambda}{\rho^2} \boldsymbol{\gamma}' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right). \end{aligned}$$

which establishes Proposition 3 in this setting where $\boldsymbol{\Omega}$ is not row-stochastic (but has spectral radius equal to one). Proposition 4 follows immediately.

We can also derive the growth rate of sectoral knowledge stock along a BGP. We have already derived the centrality-weighted growth rates above:

$$\boldsymbol{a}' \frac{\mathrm{d} \ln \boldsymbol{q}}{\mathrm{d}t} = \lambda \boldsymbol{a}' \left(\ln \boldsymbol{\eta} + \ln \bar{s} \cdot \mathbf{1} + \ln \boldsymbol{b} \right) \tag{A28}$$

We now derive the growth rate $\boldsymbol{g}$ of $\boldsymbol{q}$ in each sector. We have

$$\boldsymbol{g} = \lambda \left(\ln \boldsymbol{\eta} + \ln \bar{s} \cdot \mathbf{1} + \ln \boldsymbol{b} + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t \right)$$

Differenting with respect to time, we get

$$\boldsymbol{g} = \boldsymbol{\Omega} \boldsymbol{g}$$

Hence, the vector of sectoral growth rates is the right-Perron vector of the spillover matrix $\boldsymbol{\Omega}$, with the scale pinned down by equation (A28).

B.10 Resource Mobility Between Production and R&D

In the closed economy analysis in the main text, we assumed the endowments of production workers $\bar{\ell}$ and scientists $\bar{s}$ are both exogenous. We now argue that the optimal allocation shares $\ell_{it}/\bar{\ell}$ and $s_{it}/\bar{s}$ characterized in Lemma 1 and Proposition 1 continue to hold even if agents in the economy can endogenously choose to become workers or scientists.

First, note that the proofs of Lemma 1 and Proposition 1 continue to hold even if the exogenous endowments of workers and scientists are time-varying. Let $V\left(\boldsymbol{q}_0; \left\{\bar{\ell}_t\right\}, \left\{\bar{s}_t\right\}\right)$ denote the planner's value function, where the masses of workers and scientists are both exogenous along the entire growth path. The value function (7) in the main text corresponds to the special case where $\bar{\ell}_t = \bar{\ell}$ and $\bar{s}_t = \bar{s}$.

Now assume the economy is endowed with a unit mass of agents who can freely choose to become workers or scientists, $\bar{\ell}_t + \bar{s}_t = 1$. The value function that solves the relaxed problem, where $\bar{\ell}_t$ and $\bar{s}_t$ are endogenous, can be written as

$$V\left(\boldsymbol{q}_0\right) = \max_{\left\{\bar{\ell}_t, \bar{s}_t\right\}} V\left(\boldsymbol{q}_0; \left\{\bar{\ell}_t\right\}, \left\{\bar{s}_t\right\}\right) \quad \text{s.t. } \bar{\ell}_t + \bar{s}_t = 1.$$

Since the optimal allocation shares of workers ($\ell_{it}/\bar{\ell}$) and scientists ($s_{it}/\bar{s}$) are invariant to the total mass of workers and scientists, it follows directly that the solution characterized in Section 2.2 continues to hold in the relaxed problem.

C Details on Data Construction

In this appendix, we provide details on data collection and harmonization and robustness of our approach.

C.1 U.S. Patent Data

U.S. patent data are obtained from the United States Patent and Trademark Office (USPTO).[13] The data include information on patent inventors and patent assignee, allowing us to identify the geographic locations of the innovation (e.g., identifying cases in which a Chinese firm is granted a USPTO patent). We also observe the timing of the patents including the application and grant year. Each patent record also provides information about the invention itself, including—important for our research—its technology classifications based on the International Patent Classification (IPC) system and the citations it makes to prior inventions.

C.2 Global Patent Data

Data Source To capture global innovation, we use global patent data collected from Google Patents. The data set contains information on more than 36 million patents from the more than 40 main patent authorities around the world, over the period 1976–2020, including the USPTO, the European Patent Office (EPO), the Japanese Patent Office (JPO), and the Chinese National Intellectual Property Administration, among others. For each patent, Google Patents provides similar information as in the USPTO data described above.

Google Patents data are obtained from the DOCDB (EPO worldwide bibliographic data), the same underlying source as the more widely used PATSTAT data. We choose to use Google Patents as our main global innovation data source because it is public and accessible to all researchers free of charge. In Appendix D, we discuss specific differences between Google Patents and PATSTAT data. We show that these databases have only minor differences in their coverages and definitions of key variables and that all our empirical results are robust to both.

Identifying Patenting Locations Filing a patent in a country or patent office does not necessarily mean the underlying invention is created in the same geographic unit (e.g., Chinese firms file USPTO patents, Korean firms file patents with the Chinese National Intellectual Property Administration). These "global patenting" activities pose two important challenges for our empirical analysis. First, we need to properly determine the geographical location of the innovating activities. We assign each patent to a geographical unit according to the country of residence of its inventor(s). When a given patent is associated with multiple inventors from different countries or territories, we assign these inventors equal weight (e.g., N inventors each obtaining $1/N$ credit). If this information is not available (as in 31% of the global patent sample),[14] we use the country of

[13]We obtain the patent data from the USPTO PatentsView platform, accessible at https://www.patentsview.org/download/.

[14]Patent observations with only the country of the patent office as geographic location are mainly historical U.S. patents (51%) and historical patents originating from France, Germany, and the Soviet Union (each accounting for about 10%).

the assignee(s) instead. For 8% of patents with no easily accessible geographic location data, we assign the country of the patent office.

Identifying a Unique Invention Behind Multiple and Multinational Patents The second challenge is to de-duplicate multiple patents filed with different patent authorities for the same underlying invention. This is common practice for IP protection reasons, but may lead to double counting. To overcome this challenge, we use patent family information. We assign a set of patents to the same family if they have: (1) the same application number; or (2) the same PCT number; or (3) the same Google-provided patent family ID; or (4) at least one priority application number in common. Using patent family information, we can make sure a single invention is not counted more than once even when multiple patents are filed based on it. We also can use the earliest filing date to properly identify the timing of the underlying invention.

Cross-country Citations Importantly, patent citation information is global too—that is, we observe citations made by a patent filed by a U.S. firm with the USPTO to a patent owned by a German firm filed at the EPO. This allows us to track the innovation network at the global scale. In our sample, the proportion of citations a patent makes to foreign patents is 38%, and this number has been growing over the years.

C.3 Connecting Patent Data with Sectoral Data

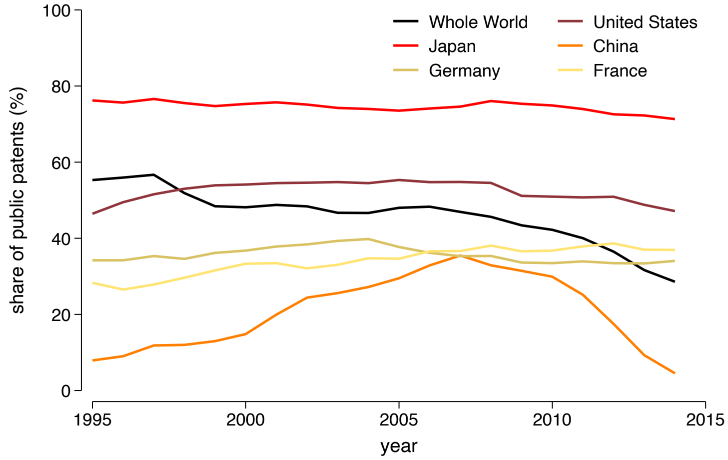
Patent data are classified into International Patent Classification (IPC) classes based on the technological content of the invention. The IPC system provides a uniform and hierarchical system of language-independent symbols for the classification of patents and utility model according to the different areas of technology to which they pertain. The IPC classification system does not naturally map to the sector classifications in either the WIOD data nor the BLS data on sectoral output and linkages. Specifically, each sector could patent in multiple IPC classes, while many sectors could patent in each single IPC class. Patent data need to be mapped to sectoral data (on value-added, R&D expenditures, employment, intermediate inputs, etc.) for our empirical analysis in different sections of our paper. This includes: (1) constructing sectoral measures of innovation activities, and (2) projecting sectoral measures into technology class levels.

Measuring Innovation at the Sector Level To construct innovation output for each country-sector-year and the country-sector-pair-wise innovation network, we need to map innovation activities to industrial sectors. We rely on our ability to observe innovation activities at the level of firms, for which we observe their industry classifications. Starting with U.S. domestic data—we link the USPTO patent database to Compustat using the bridge file provided by the NBER (up to the year 2006) and KPSS's data repository.[15] For later years, we complete the link using a fuzzy matching method based on company name, basic identity information, and innovation profiles, similar to Ma (2020) and Ma (2021). Firms' sectoral classifications are defined by North American Industry Classification System (NAICS) codes, which are then mapped to BLS sectors using the crosswalk file provided by the BLS website.[16] For each sector, we can aggregate all

[15]The extended data for KPSS can be accessed at https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data.

[16]Accessed at https://www.bls.gov/ces/naics/.

Figure A.1. Comovements of Public Patent Sample and Whole Sample



Notes. This table documents the time trend of the patent shares for firms covered in our firm-level databases across the world and in different countries.

innovation activities including patent numbers, citation-adjustment patent counts, and total R&D expenditures, conducted by U.S. firms in that specific sector..

The connection between international patent and sectoral data implements a similar logic but uses more complicated data collection and matching processes. We assemble information on global firms from Worldscope and Datastream databases accessed through Wharton Research Data Services (WRDS). The raw data sets cover more than 109,000 global firms located in 160 countries all over the world. The process is similar to that described above for U.S. data. The standard industry classifications in these databases are based on the International Standard Industrial Classification (ISIC), and can therefore be accurately mapped to the WIOD, which is also organized using the ISIC system.

The benefit of using information on firms to accurately link innovation to industrial sectors warrants the question of how representative those firms' innovation are. We find that firms in our data set produce about half of all patents in each country—for example, our sample of firms covers 44% of patents in the U.S., and 65% in Japan, two countries with the largest number of patents. Figure A.1 shows the time trend of patent shares from firms covered in our databases in the whole world and in different countries. The similarity of industry distribution between patents from covered firms and all patents in the USPTO is 0.97 when we compare the share of patents in each of the 131 3-digit IPCs for all patents and for patents from firms covered in our firm-level databases.

Projecting Sectoral Measures to Technology Classes When the unit of analysis is an IPC class (in a certain country-year), the key challenge is to project sectoral measures, such as value-added, to technology classes. We use the sector-IPC mapping provided in Lybbert and Zolas (2014). Using this mapping, we decompose each sectoral measure with proper weights to relevant IPC classes, and then aggregate the measures into the IPC level.

C.4 Constructing Cross-Sector R&D Allocation Data

Our quantitative analysis uses data on R&D allocation across different technology classes in each country. There is no standard database to exhaustively measure such information. Our primary measure relies on aggregating firm-level R&D expenditures to the country-sector-year level, based on three widely used firm-level data sets: Compustat, Worldscope, and Datastream. Combined, these data cover more than 100,000 global firms located in 160 countries and account for over 95% of the world's total market capitalization. For multinationals, we first attribute the firm-level R&D expenditures to IPC-country level in proportion to each firm's shares of patents in each IPC-country, following Griffith, Harrison, and Van Reenen (2006), and then aggregate to IPC-country-year level.

This primary measure of sectoral R&D has the advantage of covering more country-years compared to alternative approaches such as the OECD ANBERD Database. It also allows us to attribute R&D expenditures of multi-sector and multinational firms more explicitly and in a more transparent fashion. However, the primary measure of sectoral R&D is imperfect, as the firm-level data sets oversample large firms and have potentially different reporting standard across countries; we also miss R&D inputs from public sectors. Nevertheless, it is important to note that, as our theory concerns the cross-sector R&D allocation, what matters for our quantitative analysis later is the allocation shares of R&D resources across sectors in each country and not the aggregate R&D levels; any mismeasurement that affects all sectors proportionally should have no quantitative impacts.

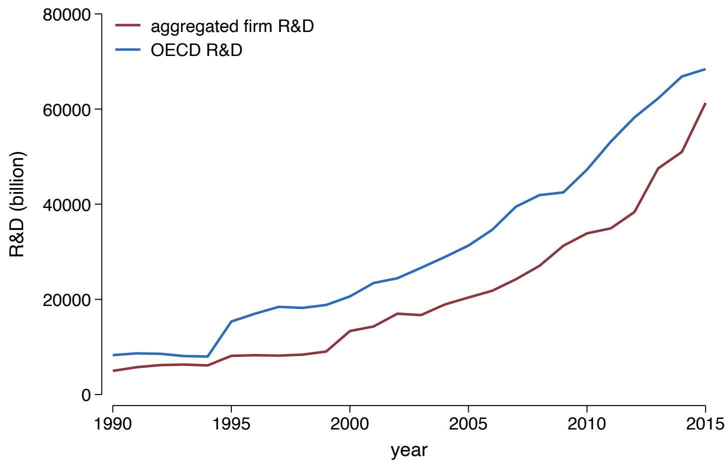
As robustness checks, we show that our primary measure of R&D allocation shares correlates strongly with two independent sources of R&D data, thereby giving us confidence in using our measure for quantitative analysis. We first provide a robustness check using the OECD Analytical Business Enterprise Research and Development (ANBERD) Database (Machin and Van Reenen, 1998), which has country-sector-level R&D information. Relative to our primary R&D measure, the ANBERD Database has more limited country-year coverage and relies more on imputations from firm-level surveys. Our primary R&D measure also allows us to explicitly and transparently attribute R&D of multi-sector or multinational enterprises to different sectors and countries.

For all the major economies in both data sets, R&D allocation from ANBERD is highly correlated with our primary measure. In the subsample of country-year observations covered in both data sources, we show that R&D expenses calculated from our firm-level data represents a significant proportion of R&D estimated by the ANBERD data, and they follow a very similar aggregate trend (Figure A.2).

The second robustness check calculates the cross-sector R&D allocation using the innovation output (which is better measured) rather than input: the number of patents produced in each country-IPC (or country-sector) divided by total number of patents produced in that specific country.

Table A.1 shows the correlation among R&D allocation measures used in our empirical analysis—R&D expense shares using R&D expenditures aggregated from firm-level data; R&D expenditures surveyed and imputed in the OECD ANBERD database; and patent shares. The correlations are calculated using 20 top patenting countries in 2010 and their R&D allocation measures across 3-digit IPC categories. The top panel first aggregate sectoral R&D expenditures across all countries and then calculate correlation of the sectoral R&D shares. The bottom panel calculate a country-specific sectoral R&D allocation correlations and then average the correlations across different

Figure A.2. Comovements of Public Patent Sample and Whole Sample



Notes. This table documents the time trend of total R&D expenditures calculated from aggregating firm-level R&D from Compustat, Worldscope, and Datastream and those calculated from aggregating country-sector information from OECD ANBERD data. For each year, we cover countries that are covered in both databases.

countries. In each panel, the bottom half of the table shows the Pearson correlations; the top half of the table shows Spearman's rank correlation, which is equal to the Pearson correlation of the rank values.

These three proxies for R&D allocations are highly correlated. For example, in Panel A, the correlation between R&D allocations aggregated from firm-level data and from the OECD scores above 0.9. The correlation between input shares and the patent output shares is slightly lower, but still above 0.8. The high correlations among these three measures of R&D allocation shares translate into the robustness of our quantitative results, as illustrated in Section E.4 of the Online Appendix.

D Cross-checking Google Patents with PATSTAT

This appendix compares data from Google Patents (accessible to all researchers free of charge) and the widely used commercial database PATSTAT. These exercises will compare their data coverage, key variable definitions, and the robustness of empirical analyses in those two databases.

D.1 Basic Data Structure and Coverage

Google Patents and PATSTAT share nearly identical data structure. Both databases have three levels of innovation units: publication, application, and family.

- **Application:** The central unit is an innovation application, which is a request filed to a patent office for patent protection for an invention (which may or may not be granted later).

Table A.1. Different Measures of Cross-Sector R&D Allocation Are Highly Correlated

Panel A	Share of Aggregated Firm R&D	Share of Patents	Share of OECD R&D
Share of Aggregated Firm R&D		0.83	0.97
Share of Patents	0.86		0.82
Share of OECD R&D	0.93	0.78	

Panel B	Share of Aggregated Firm R&D	Share of Patents	Share of OECD R&D
Share of Aggregated Firm R&D		0.74	0.91
Share of Patents	0.74		0.76
Share of OECD R&D	0.74	0.69	

Notes: This table shows the correlation of R&D allocation measures used in our empirical analysis–R&D expense shares using R&D expenditures aggregated from firm-level data; R&D expenditures surveyed and imputed in the OECD ANBERD database; and patent shares. The correlations are calculated using 20 top patenting countries in 2010 and their R&D allocation measures across 3-digit IPC categories. The top panel first aggregate sectoral R&D expenditures across all countries and then calculate correlation of the sectoral R&D shares. The bottom panel calculate a country-specific sectoral R&D allocation correlations and then average the correlations across different countries. In each panel, the bottom half of the table shows the Pearson correlations; the top half of the table shows Spearman's rank correlation, which is equal to the Pearson correlation of the rank values.

- **Publication (most basic unit):** After an application is filed, various publications could be issued.[17] These publications can be disclosed patent filings (often 18 months after the initial filing date), granted patent specification, corrections, etc. In simple terms, publications help identify key events over an application's life cycle. The basic units of both Google Patents and PATSTAT are innovation "publications."

- **Family:**[18] Applications that cover the same underlying invention are grouped into families. This often happens when the same invention is filed with multiple patent offices, sometimes simultaneously, for protections in different countries. All applications (and publications tracking their life cycle events) in the same family thus have the same priorities, and their technical content is often regarded as identical or almost identical. Patent family counting allows us to track unique inventions across different economies.

Figure A.3 presents the sample coverage of publications, the most basic units, for both Google Patents and PATSTAT in the time series. The coverages of the two data sets are virtually identical.

D.2 Identifying Granted Patents

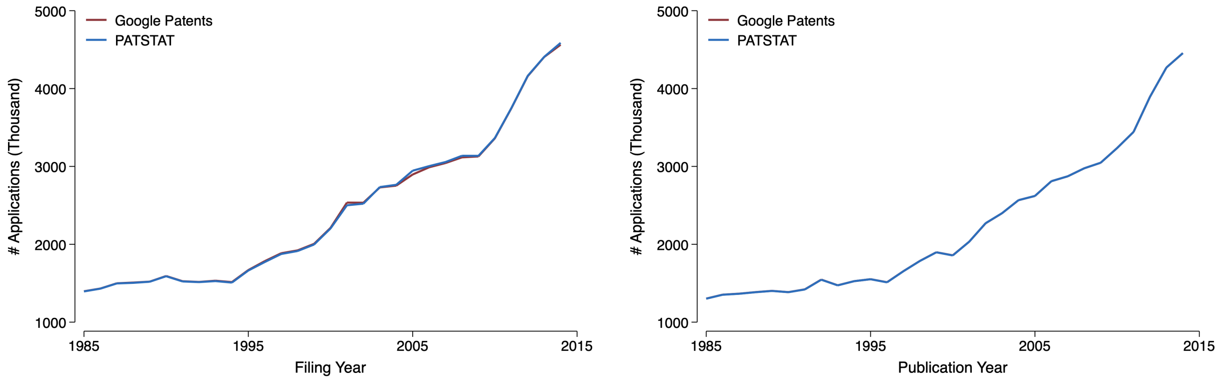
Publications represent the most comprehensive set of innovation-related documents, yet many of them are irrelevant for studying innovation—some publications are associated with denied applications, some are design patents unrelated to scientific or technological progress, etc. As a result, it is useful to identify granted patents related to new technologies (e.g., utility patents in the USPTO

[17]In cases that generate no publications (i.e., the invention is treated with absolute confidentiality), the invention would not be accessible in any database.

[18]In our paper, we consider the more widely accepted definition of simple family, also called the DOCDB family or Espacenet patent family.

Figure A.3. Google Patents v.s. PATSTAT by Year

(a) Applications by Filing Year (b) Applications by Publication Year



system). The two database handle this process largely identically, yielding very comparable patent sets. However, there are three noticeable differences:

1. Identifying whether a patent is granted mainly relies on the kind code of the patent, which is defined by the patent office and will change with the reform of the patent system of the patent office.[19] For example, the kind code of patent "US-10001017-B2" is "B2." The rules used to identify granted patents differ somewhat in Google Patents vs. PATSTAT.

2. Because PATSTAT uses additional legal event data to identify granted patents, patents granted by some small patent offices can be identified.

3. Other minor differences include missing filing dates or issue dates.

Table A.2 shows the comparison of granted patents between Google Patents and PATSTAT and list the sources of coverage differences.

[19] For the detailed meaning of difference kind codes in different patent offices, we refer readers to the document of format concordance of publication numbers in EPO (see https://www.epo.org/searching-for-patents/data/coverage/regular.html).

Table A.2. Difference of Granted Patents Between Google Patents and PATSTAT

Panel (A): For Granted Patents in PATSTAT

	# Patents			
Patents granted in 1985—2014	19,923,292	100.00%		
Overlapped with Google Patents	17,135,611	86.01%		
Non-overlapped with Google Patents	2,787,681	13.99%	100.00%	
1. Additional patent office data from legal event data	1,456,242		52.24%	100.00%
(1) For patent office ZA	175,317			12.04%
(2) For patent office MX	125,298			8.60%
(3) For patent office PL	125,246			8.60%
(4) For patent office UA	95,956			6.59%
(5) For patent office PT	82,533			5.67%
(6) For patent office DD	79,171			5.44%
(7) For patent office NO	65,312			4.48%
(8) For patent office BR	62,447			4.29%
(9) For patent office HU	61,707			4.24%
(10) For patent office IL	57,165			3.93%
(11) Other patent offices including BG, BY, CH, CO, CS, CU, CZ, EA, EE, GE, GR, HK, HR, ID, IE, IN, IS, KE, LT, LV, MA, MC, MD, ME, MN, MT, MY, NI, OA, PE, PH, RO, RS, SA, SE, SG, SI, SK, SM, SV, TJ, TR, UY, VN, YU, ZW	526,090			36.13%
2. Additional rules used to identify granted patents	1,331,439		47.76%	100.00%
(1) For patent office AT, patents with kind code in [T]	543,805			40.84%
(2) For patent office DE, patents with kind code in [T2]	468,202			35.17%
(3) For patent office KR, patents with kind code in [A]	65,237			4.90%
(4) For patent office DK, patents with kind code in [T3]	58,520			4.40%
(5) For patent office ES, patents with kind code in [A1, A6]	47,354			3.56%
(6) For patent office AU, patents with kind code in [A1, A8]	32,835			2.47%
(7) For patent office FI, patents with kind code in [C]	31,907			2.40%
(8) For patent office CN, patents with kind code in [A]	28,928			2.17%
(9) For patent office AR, patents with kind code in [A1]	24,865			1.87%
(10) For patent office US, patents with kind code in [E]	16,366			1.23%
(11) Other patent offices	13,420			1.01%

Panel (B): For Granted Patents in Google Patents

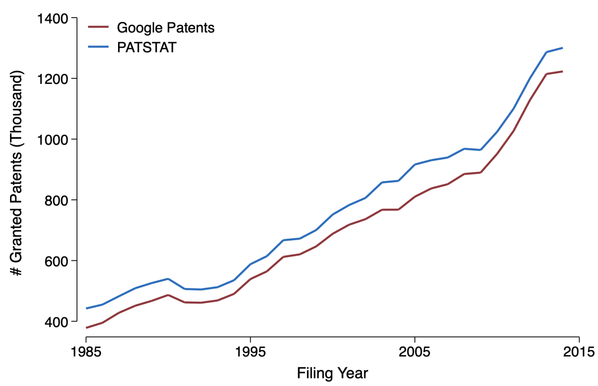
	# Patents			
Patents granted in 1985—2014	18,144,529	100.00%		
Overlapped with PATSTAT	17,135,612	94.44%		
Non-overlapped with PATSTAT	1,008,917	5.56%	100.00%	
1. Additional patent office data from legal event data	0		0.00%	
2. Additional rules used to identify granted patents	1,008,917		100.00%	100.00%
(1) For patent office DE, patents with kind code in [D1]	883,482			87.57%
(2) For patent office DK, patents with kind code in [T3]	58,118			5.76%
(3) For patent office FI, patents with kind code in [B]	31,585			3.13%
(4) For patent office BE, patents with kind code in [A3, A4, A5, A6, A7]	20,797			2.06%
(5) For patent office KR, patents with kind code in [B1]	6,546			0.65%
(6) For patent office ES, patents with kind code in [B1]	2,399			0.24%
(7) For patent office DZ, patents with kind code in [A1]	1,755			0.17%
(8) For patent office AU, patents with kind code in [B2]	1,458			0.14%
(9) For patent office EP, patents with kind code in [B1]	1,344			0.13%
(10) For patent office SU, patents with kind code in [A1]	932			0.09%
(11) Other patent offices	501			0.05%

Notes. This table compares coverages of granted patents between Google Patents and PATSTAT and the reasons for discrepancies.

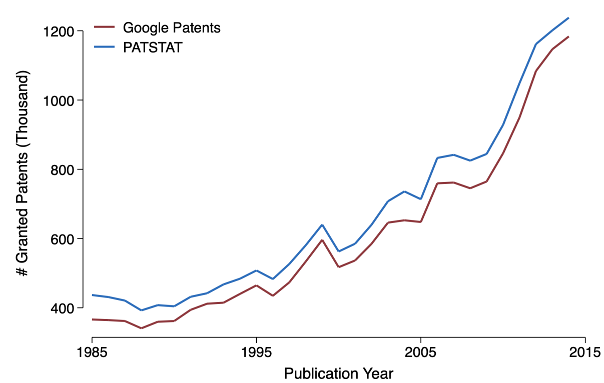
Despite those differences, Google Patents and PATSTAT agree on roughly 95% of the identified granted patents. In Figure A.4, we present the numbers of granted patents in Google Patents and PATSTAT. We also show this difference across various patent offices and countries of origin.

Figure A.4. Google Patents v.s. PATSTAT Coverage

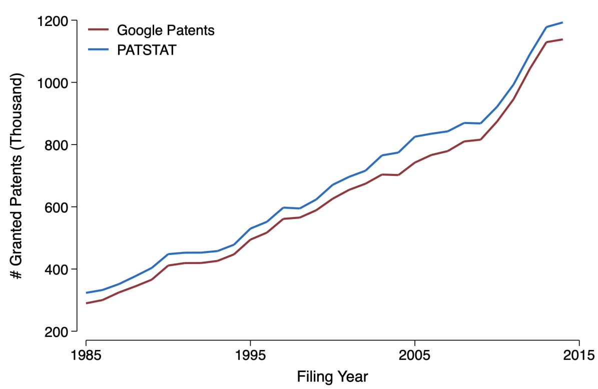
(a) Granted Patents by Filing Year



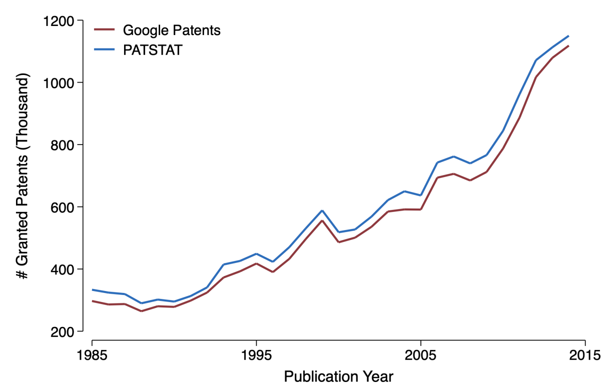
(b) Granted Patents by Publication Year



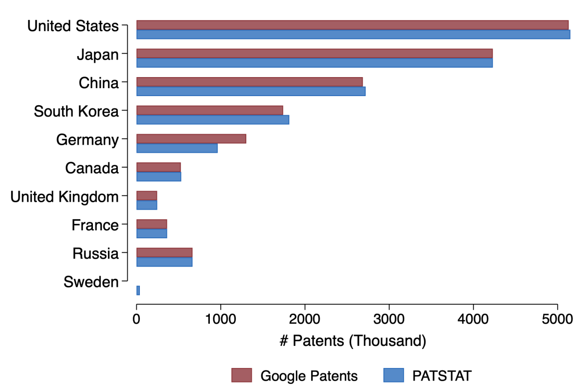
(c) Granted Patents by Filing Year (43 WIOD Countries)



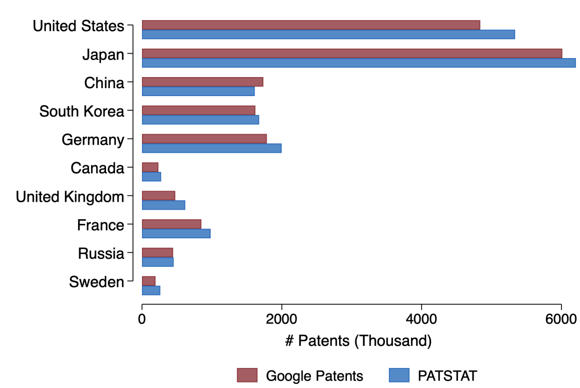
(d) Granted Patents by Publication Year (43 WIOD Countries)



(e) Google Patents vs. PATSTAT By Patent Offices



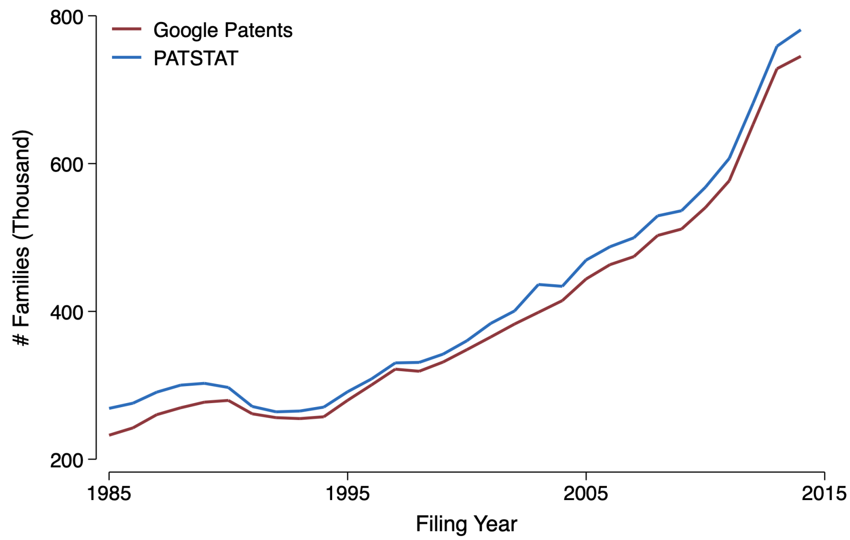
(f) Google Patents vs. PATSTAT By Invention Origin



D.3 Patent Family

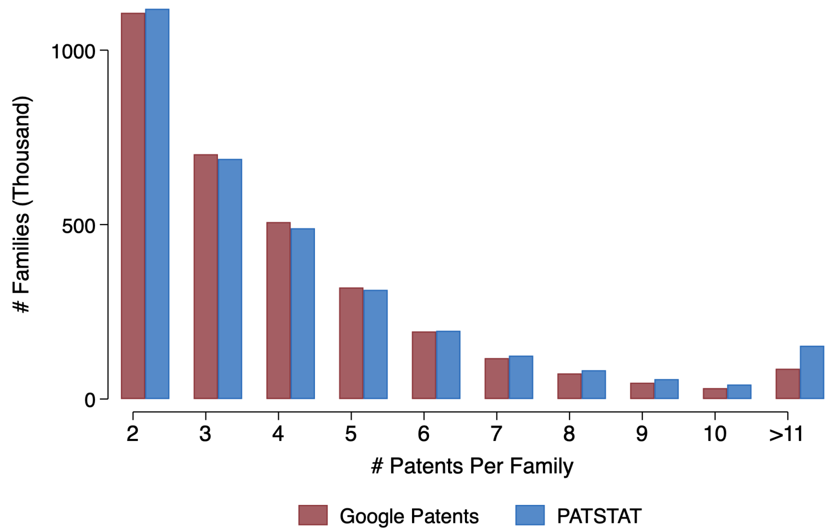
Defining patent family involves the use of information regarding priority dates and priority patents in the global patent database, among others. Figure A.5 presents the number of patent families identified in both data sets. They are very comparable to each other, and the minor gap can be explained by the differences in the number of identified patents described in the previous section.

Figure A.5. Google Patents v.s. PATSTAT: Patent Families by Year



To further check this consistency, in Figure A.6 we show the distribution of the number of patents in each family in Google Patents and PATSTAT, which again are quite comparable. In Google Patents, there are 11,693,980 patent families between 1985 and 2014. Among these families, 3,184,884 contain at least two patents, and on average, these families contain 3.99 patents. In PATSTAT, there are 12,344,446 patent families between 1985 and 2014. Among those families, 3,263,376 of them contain at least two patents, and on average, these families contain 4.34 patents.

Figure A.6. Google Patents v.s. PATSTAT: Distribution of Number of Patents in Each Family



We next perform a family-to-family comparison between the two databases. First, we focus on families that only contain one patent: 98.74% of these families in Google Patents are consistent with that in PATSTAT, and 97.79% of those families in PATSTAT are consistent with those in Google Patents. For patent families with two patents, the share of patents in PATSTAT that are consistent with Google Patents is 94.11%; the share of patents in Google Patents that is consistent with PATSTAT is 94.38%. Overall, patent families seem to be consistently defined across the databases at a very high rate.

D.4 Robustness of Results Using Google Patents and PATSTAT

In this section, we present results from using PATSTAT patent data as the base for innovation measurement and innovation network construction. The overall takeaway is that the results using PATSTAT are virtually identical to results using Google Patents.

D.4.1 Innovation Network

Results in this subsection show that innovation networks constructed using PATSTAT and Google Patents are highly correlated (Table A.3), and they have virtually identical properties such as centrality (Figure A.7) and visualizations (Figure A.8).

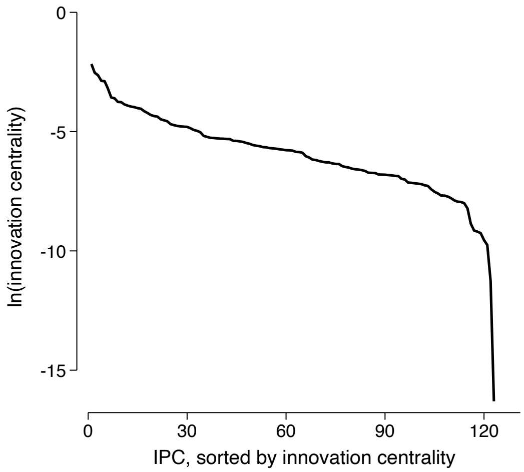
Table A.3. Correlations of Between the Innovation Network from Google Patents and PATSTAT

All	U.S.	Japan	China	Korea	Germany	Canada	UK	France	Russia	Sweden
0.997	0.998	0.945	0.987	0.975	0.979	0.986	0.989	0.966	0.887	0.934

Notes. This is the correlation between the innovation networks calculated using Google Patents and PATSTAT data.

Figure A.7. Innovation Centrality and Key Sectors for PASTAT

(a) Innovation Centrality Across IPCs



(b) Top Ten IPCs by Innovation Centrality a_i

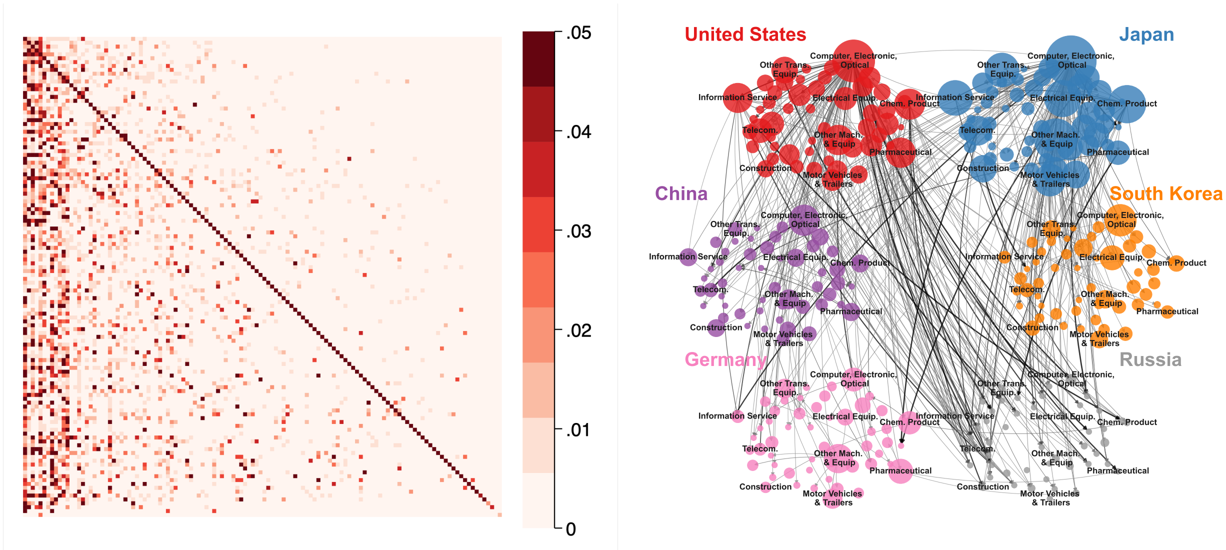
1	A61	medical or veterinary science; hygiene
2	G06	computing; calculating or counting
3	H01	basic electric elements
4	G01	measuring; testing
5	H04	electric communication technique
6	B60	vehicles in general
7	G02	optics
8	B01	physical or chemical processes or apparatus in general
9	C08	organic macromolecular compounds; their preparation or chemical working-up; compositions based thereon
10	F16	engineering elements or units; general measures for producing and maintaining effective functioning of machines or installations; thermal insulation in general

Notes. This figure reproduces Figure 2 in the paper using PATSTAT data. This figure presents the innovation centrality of different technology classes categorized using IPCs. Panel (a) plots $\log(a_i)$, and the sectors are ranked in descending order based on a_i. Panel (b) lists the top ten IPCs by their innovation centrality.

Figure A.8. Visualizing the Innovation Network for PATSTAT

(a) IPC-to-IPC (131×131) Network Ω

(b) Global Innovation Network Across Country-Sectors



Notes. This figure reproduces Figure 1 in the paper using PATSTAT data. The left panel visualizes the IPC-to-IPC network **Ω** as a heatmap, with darker colors representing larger matrix entries; sectors are ordered according to their innovation centrality. The right panel visualizes the global innovation network. Each node is a country-sector, with size drawn in proportion to patent output. Arrows represent knowledge flows, with width drawn in proportion to citation shares.

D.4.2 Knowledge Spillovers

This subsection reproduces results to confirm the mechanism of sectoral innovation activities being influenced by innovation from global upstream sectors.

Table A.4. Evidence of the Global Innovation Network for Knowledge Spillovers
Based on WIOD - PATSTAT

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge^{Up}_{mit}$	0.194***	0.203***	0.185***	0.308***	0.335***	0.300***
	(0.050)	(0.052)	(0.051)	(0.075)	(0.075)	(0.076)
$ln(R\&D\ Stock)_{mi,t-1}$	0.041***	0.041***	0.041***	0.076***	0.076***	0.075***
	(0.013)	(0.013)	(0.013)	(0.018)	(0.018)	(0.019)
$Knowledge^{Down}_{mit}$		-0.025			-0.078**	
		(0.031)			(0.039)	
$Knowledge^{Up,IO}_{mit}$			0.044			-0.068
			(0.067)			(0.068)
R^2	0.967	0.967	0.968	0.942	0.942	0.942
No. of Country x Sectors	570	570	556	570	570	556
No. of Obs	11011	11011	10771	11011	11011	10771
Fixed Effects	Country x Sector, Country x Year, Sector x Year					

Notes. This table reproduces Table 3 in the paper using PATSTAT data. This table tests the relation between innovation in a focal sector and past innovation in connected sectors through the innovation network, in an international setting. We restrict the sample to country-sectors that have at least ten patents over the full sample period. To measure innovation production (Y), we use the number of patents and total number of citations. The key variable of interest, $\text{Knowledge}^{Up}_{it}$, is the knowledge from upstream, defined in (28). Fixed effects at the country-sector, country-year, and sector-year levels are included as controls. Columns (2) and (5) include downstream knowledge as a control. Columns (3) and (6) include knowledge accumulated from upstream sectors in the production network as a control. Standard errors in parentheses are clustered at the country-sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

E Supplementary Results

In this section, we provide additional empirical results.

E.1 Innovation Networks Are Stable Over Time and Across Countries

We first document that innovation networks are stable over time and across innovative countries. We construct time-varying measures of the innovation network, following the formula in (24) but using citations made by patents filed during specific time periods, from all countries in our sample. For the innovation network time-stamped at t, we use new patents and their citations between $t-10$ and $t-1$ to construct the network. Table A.5 shows the correlations between our baseline, time-invariant measure ω_{ij} of the innovation network and these other measures ω_{ijt} constructed using patents filed in specific years t. The bottom half of the table shows the Pearson correlations; the top half of the table shows Spearman's rank correlation, which is equal to the Pearson correlation of the rank values and can be more revealing of network similarities than the Pearson correlation of values (Liu, 2019). Table A.5 shows that the innovation network is highly stable over time; the time-varying measures exhibit above 0.8 correlations even when measured using citation data that are three decades apart, and all year-specific measures are strongly correlated with our time-invariant baseline measure.

Table A.5. The Innovation Network Is Highly Correlated Over Time

Time Period	All years	2020	2010	2000	1990	1980
All years		0.98	0.98	0.97	0.90	0.89
2020	0.95		0.97	0.93	0.86	0.85
2010	0.96	0.97		0.96	0.88	0.87
2000	0.93	0.92	0.96		0.92	0.90
1990	0.90	0.80	0.84	0.90		0.91
1980	0.81	0.77	0.81	0.87	0.89	

Notes: This table shows the correlation of innovation networks calculated using different vintages of patent data. For each decade, all global patents in that decade are included when constructing the innovation network. The bottom half of the table shows the Pearson correlations; the top half of the table shows Spearman's rank correlation, which is equal to the Pearson correlation of the rank values.

Second, we construct country-specific innovation networks. Specifically, we use the same formula (24) but restrict the sample to all patents from each country. Table A.6 shows the correlations between our baseline, location-invariant measure and the country-specific measures for the ten countries with the most patents in our sample; Pearson correlations are again shown in the bottom half of the table whereas Spearman's rank correlations are shown in the top half. Innovation networks are highly stable across countries. In particular, our baseline measure, which is constructed using patents pooled from around the world, has a correlation coefficient of above 0.98 with the network implied by U.S. patents and is also highly correlated (>0.8 rank correlation) with the innovation networks in Japan, China, Germany, Canada, the U.K., and France. The only exception is Russia, whose innovation network is less perfectly correlated with the measures, but the correlation is still substantial (about 0.6).

Table A.6. The Innovation Network Is Highly Correlated Across Countries

Countries	All	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
All		0.98	0.87	0.87	0.84	0.89	0.63	0.86	0.92	0.88	0.81
US	0.95		0.84	0.86	0.82	0.88	0.64	0.85	0.92	0.88	0.80
Japan	0.86	0.83		0.88	0.89	0.85	0.63	0.87	0.86	0.84	0.83
China	0.85	0.86	0.87		0.88	0.85	0.66	0.85	0.87	0.86	0.82
South Korea	0.78	0.77	0.83	0.84		0.84	0.64	0.84	0.85	0.82	0.84
Germany	0.85	0.87	0.81	0.80	0.72		0.64	0.83	0.87	0.83	0.81
Russia	0.62	0.63	0.62	0.62	0.55	0.61		0.65	0.64	0.64	0.66
France	0.91	0.86	0.79	0.77	0.72	0.82	0.57		0.86	0.85	0.83
UK	0.87	0.89	0.85	0.85	0.80	0.86	0.64	0.80		0.88	0.82
Canada	0.86	0.88	0.79	0.81	0.71	0.81	0.59	0.80	0.81		0.81
Netherlands	0.84	0.85	0.79	0.82	0.75	0.79	0.58	0.78	0.79	0.81	

Notes: This table shows the correlation of innovation networks calculated using patents in the top ten innovative countries ranked by patent outputs between 2010—2014. When calculating this country-specific innovation network, all patents of the country across all years are included. The bottom half of the table shows the Pearson correlations; the top half of the table shows Spearman's rank correlations, which are equal to the Pearson correlation of the rank values.

E.2 Knowledge Spillovers Through Innovation Networks—Robustness

This subsection provides additional robustness analyses on innovation diffusion through innovation networks, echoing Section 4.2 in the paper. The main results supporting the important role of innovation networks in knowledge spillovers are provided in Tables 2 and 3 in the paper. Below, we present tests to show the robustness of these results. Specifically, these analyses incorporate changing U.S. BLS Sectors to IPC (International Patent Classification) classes as the node in innovation networks (Table A.7), additional measures of innovation output (Table A.11), and different time horizons to calculate upstream innovation (Tables A.9 and A.12). Finally, we revisit the dynamic prediction of our key law of motion (25), that upstream knowledge from the more distant past has less effect on patent output, in Figure (A.9). The figure shows an obsolescence-like pattern (Ma, 2021) in which past upstream knowledge's effect on subsequent innovation weakens over time, precisely as our theory predicts.

Table A.7. U.S. and Global Evidence of Knowledge Spillover Through Innovation Networks
Based on IPC

	US		Global	
$Y=$	ln(Patents)	ln(Cites)	ln(Patents)	ln(Cites)
	(1)	(2)	(3)	(4)
$Knowledge_{it}^{Up}$	0.499***	0.523***	0.043***	0.074***
	(0.085)	(0.106)	(0.010)	(0.014)
$ln(R\&D\ Stock)_{i,t-1}$	0.409***	0.495***	-0.002	0.002
	(0.110)	(0.140)	(0.005)	(0.007)
$Knowledge_{it}^{Down}$	-0.244***	-0.349***	-0.031***	-0.025**
	(0.064)	(0.090)	(0.008)	(0.011)
R^2	0.960	0.948	0.945	0.902
No. of Sectors	431	431		
No. of Country x Sectors			4595	4595
No. of Obs	8620	8620	86224	86224
Fixed Effects	Sector, Year		Country x Sector Country x Year Sector x Year	

Notes. This table reproduces Tables 2 and 3 in the paper. The key difference is this table uses the country by detailed 4-digit IPC (international patent classification) class as the unit of nodes instead of country by (BLS or WIOD) industrial sectors.

Table A.8. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Adding The Impact of Own Sector

$Y=$	ln(Patents)			ln(Cites)			ln(Patent Value)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Knowledge_{it}^{Up}$	0.759***	0.827***	0.706***	0.947***	0.849***	0.900***	0.986***	1.043***	0.970***
	(0.131)	(0.172)	(0.130)	(0.191)	(0.270)	(0.196)	(0.293)	(0.382)	(0.286)
$Knowledge_{it}^{Own}$	0.643***	0.630***	0.577***	0.559***	0.577***	0.498***	0.252**	0.242**	0.214
	(0.054)	(0.053)	(0.080)	(0.085)	(0.080)	(0.127)	(0.105)	(0.105)	(0.184)
$ln(R\&D\ Stock)_{i,t-1}$	0.177**	0.178**	0.164**	0.130	0.128	0.117	0.412***	0.412***	0.403**
	(0.073)	(0.072)	(0.076)	(0.105)	(0.105)	(0.108)	(0.152)	(0.152)	(0.154)
$Knowledge_{it}^{Up,IO}$			-0.038			-0.054			-0.145
			(0.121)			(0.185)			(0.212)
$Knowledge_{it}^{Own,IO}$			0.092			0.088			0.071
			(0.083)			(0.108)			(0.178)
$Knowledge_{it}^{Down}$		-0.125			0.182			-0.105	
		(0.143)			(0.270)			(0.304)	
R^2	0.937	0.937	0.937	0.910	0.910	0.910	0.888	0.888	0.888
No. of Sectors	95	95	95	95	95	95	95	95	95
No. of Obs	1892	1892	1892	1892	1892	1892	1892	1892	1892
Fixed Effects		Sector, Year			Sector, Year			Sector, Year	

Notes. This table reproduces Table 2 in the paper by incorporating patenting activities from past innovation from own sector.

Table A.9. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Different Knowledge Periods

Panel (A): $\tau = 5$

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{it}^{Up,\tau=5}$	0.450***	0.513***	0.415***	0.697***	0.763***	0.670***
	(0.142)	(0.161)	(0.139)	(0.157)	(0.163)	(0.158)
$ln(R\&D\ Stock)_{i,t-1}$	0.449***	0.463***	0.430***	0.372***	0.386***	0.356***
	(0.096)	(0.096)	(0.092)	(0.112)	(0.112)	(0.109)
$Knowledge_{it}^{Down,\tau=5}$		-0.142			-0.148	
		(0.157)			(0.095)	
$Knowledge_{it}^{Up,IO}$			0.277*			0.218
			(0.164)			(0.202)
R^2	0.916	0.916	0.917	0.900	0.900	0.900
No. of Sectors	95	95	95	95	95	95
No. of Obs	1900	1900	1900	1900	1900	1900
Fixed Effects		Sector, Year			Sector, Year	

Panel (B): $\tau = 20$

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{it}^{Up,\tau=20}$	0.644***	0.694***	0.592***	0.858***	0.912***	0.819***
	(0.185)	(0.202)	(0.180)	(0.207)	(0.215)	(0.200)
$ln(R\&D\ Stock)_{i,t-1}$	0.410***	0.418***	0.397***	0.323***	0.332***	0.313***
	(0.099)	(0.100)	(0.095)	(0.113)	(0.113)	(0.110)
$Knowledge_{it}^{Down,\tau=20}$		-0.123			-0.133	
		(0.164)			(0.103)	
$Knowledge_{it}^{Up,IO}$			0.238			0.178
			(0.165)			(0.201)
R^2	0.917	0.917	0.918	0.900	0.900	0.900
No. of Sectors	95	95	95	95	95	95
No. of Obs	1900	1900	1900	1900	1900	1900
Fixed Effects		Sector, Year			Sector, Year	

Notes. This table reproduces Table 2 in the paper. The key difference is using different τ periods to calculate knowledge accumulated through the innovation network. Table 2 uses $\tau = 10$, while this table uses alternative values of $\tau = 5$ and $\tau = 10$.

Table A.10. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Exponential Knowledge Discounting

$Y=$	ln(Patents)			ln(Cites)			ln(Patent Value)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Knowledge_{it}^{Up,depreciation}$	0.507***	0.563***	0.466***	0.766***	0.824***	0.734***	0.913***	0.952***	0.916***
	(0.168)	(0.189)	(0.163)	(0.181)	(0.190)	(0.177)	(0.282)	(0.292)	(0.281)
$ln(R\&D\ Stock)_{i,t-1}$	0.431***	0.439***	0.414***	0.344***	0.353***	0.331***	0.503***	0.509***	0.504***
	(0.100)	(0.101)	(0.096)	(0.115)	(0.115)	(0.112)	(0.146)	(0.148)	(0.145)
$Knowledge_{it}^{Down,depreciation}$		-0.123			-0.127			-0.085	
		(0.159)			(0.097)			(0.117)	
$Knowledge_{it}^{Up,IO}$			0.271			0.211			-0.015
			(0.165)			(0.203)			(0.206)
R^2	0.916	0.916	0.917	0.900	0.900	0.900	0.886	0.886	0.886
No. of Sectors	95	95	95	95	95	95	95	95	95
No. of Obs	1900	1900	1900	1900	1900	1900	1900	1900	1900
Fixed Effects		Sector, Year			Sector, Year			Sector, Year	

Notes. This table reproduces Table 2 in the paper by incorporating exponential discounting of knowledge stocks.

Table A.11. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Additional Innovation Measure

$Y=$	ln(Patents)			
	(1)	(2)	(3)	(4)
$Knowledge_{it}^{Up}$	0.909***	0.945***	0.914***	1.435***
	(0.294)	(0.306)	(0.294)	(0.437)
$ln(R\&D\ Stock)_{i,t-1}$	0.500***	0.505***	0.502***	0.471**
	(0.146)	(0.148)	(0.146)	(0.188)
$Knowledge_{it}^{Down}$		-0.078		
		(0.114)		
$Knowledge_{it}^{Up,IO}$			-0.027	
			(0.208)	
$ln(R\&DTaxPrice)_{mi,t-1}$				9.757
				(9.215)
Specification	OLS	OLS	OLS	IV 2nd Stage
IV 1st Stage F-statistics				427
R^2	0.886	0.886	0.886	0.112
No. of Sectors	95	95	95	95
No. of Obs	1900	1900	1900	1140
Fixed Effects		Sector, Year		

Notes. This table reproduces Table 2 in the paper with the additional innovation measure of patent value from (Kogan et al., 2017) based on the stock market reaction to patent approval.

Table A.12. Global Evidence of Knowledge Spillover Through Innovation Networks
Different Knowledge Periods

Panel (A): $\tau = 5$

Y=	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{mit}^{Up,\tau=5}$	0.101**	0.111**	0.092*	0.223***	0.243***	0.219***
	(0.047)	(0.049)	(0.047)	(0.070)	(0.075)	(0.071)
$ln(R\&D\ Stock)_{i,t-1}$	0.044***	0.044***	0.044***	0.085***	0.085***	0.084***
	(0.013)	(0.013)	(0.013)	(0.018)	(0.018)	(0.018)
$Knowledge_{mit}^{Down,\tau=5}$		-0.017			-0.035	
		(0.032)			(0.046)	
$Knowledge_{mit}^{Up,IO}$			0.070			-0.055
			(0.065)			(0.068)
R^2	0.968	0.968	0.968	0.942	0.943	0.943
No. of Country x Sectors	570	570	556	570	570	556
No. of Obs	11014	11014	10774	11014	11014	10774
Fixed Effects	Country x Sector, Country x Year, Sector x Year					

Panel (B): $\tau = 20$

Y=	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{mit}^{Up,\tau=20}$	0.204***	0.232***	0.201***	0.425***	0.472***	0.424***
	(0.057)	(0.059)	(0.058)	(0.080)	(0.083)	(0.081)
$ln(R\&D\ Stock)_{i,t-1}$	0.043***	0.044***	0.043***	0.084***	0.084***	0.083***
	(0.013)	(0.013)	(0.013)	(0.018)	(0.018)	(0.018)
$Knowledge_{mit}^{Down,\tau=20}$		-0.079*			-0.134*	
		(0.045)			(0.069)	
$Knowledge_{mit}^{Up,IO}$			0.072			-0.051
			(0.064)			(0.068)
R^2	0.968	0.968	0.968	0.943	0.943	0.943
No. of Country x Sectors	570	570	556	570	570	556
No. of Obs	11014	11014	10774	11014	11014	10774
Fixed Effects	Country x Sector, Country x Year, Sector x Year					

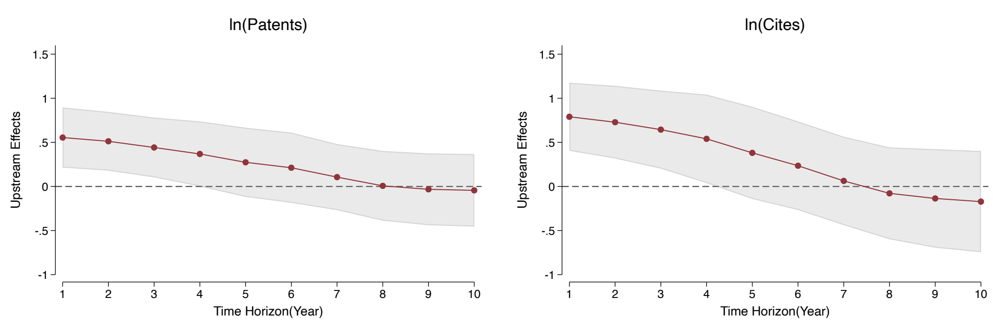
Notes. This table reproduces Table 3 in the paper. The key difference is this table uses different τ periods to calculate knowledge accumulated through the innovation network. Table 3 uses $\tau = 10$, while this table uses alternative values of $\tau = 5$ and $\tau = 10$.

Table A.13. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Exploring the I-O Linkages

Y=	ln(Patents)		ln(Cites)		ln(Patent Value)	
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{it}^{Up}$	0.509***		0.756***		0.914***	
	(0.169)		(0.192)		(0.294)	
$ln(R\&D\ Stock)_{i,t-1}$	0.410***	0.442***	0.328***	0.375***	0.502***	0.559***
	(0.096)	(0.094)	(0.111)	(0.107)	(0.146)	(0.148)
$Knowledge_{it}^{Down}$						
$Knowledge_{it}^{Up,IO}$	0.258	0.338**	0.198	0.316	-0.027	0.117
	(0.165)	(0.166)	(0.203)	(0.205)	(0.208)	(0.218)
R^2	0.917	0.914	0.900	0.896	0.886	0.881
No. of Sectors	95	95	95	95	95	95
No. of Obs	1900	1900	1900	1900	1900	1900
Fixed Effects	Sector, Year		Sector, Year		Sector, Year	

Notes. This table reproduces Table 2 in the paper by incorporating standalone knowledge spillovers from the I-O network in columns (2), (4), and (6).

Figure A.9. Dynamic Responses of Innovation Output to Upstream Knowledge



Notes. This figure presents how the focal sector's innovations dynamically respond to past innovations from upstream sectors in the innovation network. The coefficients are from regressions of focal sectors' innovations at times $t+1$ through $t+10$ on upstream knowledge measured at time-t. We control for log R&D with time-1 lag as well as sector and year fixed effects. The half-life of the dynamic effects is about 4 years.

E.3 Using R&D Tax Credit as an Instrument for Upstream R&D

Our analysis on the impact of upstream innovation (i.e., Tables 2 and 3) is subject to the concern of common shocks: a group of sectors connected to each other via citation linkages may face similar demand, supply, and investment opportunities, leading to co-movements of innovation activities. Serial correlations in these common shocks would lead to a positive coefficient β_1 in regression (27) even without cross-sector knowledge spillovers. This is a classic version of the "reflection problem" documented in Manski (1993) and, more relevant to our setting, in Bloom et al. (2013). As noted in Bloom et al. (2013), since knowledge spillovers through the innovation network are entered lagged at least one year (and up to ten years), and because fixed effects and other controls are included in the estimation, the potential bias is likely small. Nevertheless, to further resolve this issue, we consider an instrumental variable strategy based on R&D tax credits, a method widely used in innovation literature. Here we present only the basic framework and how we adapt the strategy to our setting. We refer readers to a classic use case in Bloom et al. (2013) and the Online Appendix of the paper.

This instrumental variable strategy shocks R&D activities using the user cost of R&D capital, which in turn is often closely tied to tax policies and subsidies like R&D tax credit. User cost of R&D is affected by two types of R&D tax credit, federal tax rules that interact with different firms differently (e.g., based on past R&D expenses, etc.), and state-level tax credits, depreciation allowances, and corporation taxes that affect firms differently based on the location of R&D activities.

- For state-level tax credits, we obtain the state-by-year R&D tax price data, available for 1970 to 2006, from Wilson (2009). These data are further aggregated to sector-year-level tax price of R&D by calculating the weighted sectoral average, which is weighted using the total number of inventors in a sector who work in each state (ten-year average of inventor shares). In other words, if a sector has more inventor weight in a high tax credit state (thus the user cost is lower), the sector will have a lower user cost of R&D in our aggregation. Using inventor shares is common practice in this literature as R&D labor cost is often the key target of R&D tax policies.

- For the federal tax component, which is shown to be less powerful for explaining sector-level R&D activities in our setting, we follow the approach in Bloom et al. (2013) and construct a firm-year level federal tax-driven user cost of R&D. This firm-year-level measure is then further aggregated to sector-year level by weighting each firm according to its size measured using the number of inventors.

The R&D user cost can also be calculated at the country-sector-year level. For this purpose, we obtain data from Thomson (2017), who provides the user cost estimates for different types of R&D input, in particular labor and capital, in different country-years. Following Thomson (2017), we calculate the tax price at the country-sector-year level using the weight-average tax price of different expenditure types with lagged expenditure share on those types as weights. For example, the "Apparel, dressing, and dyeing of fur" industry has a capital-labor R&D composition ratio of 92% to 8%, then the R&D user cost is a weighted average using those ratios. This estimate covers 25 WIOD countries from 1980 to 2006.

We implement the empirical strategy by first projecting sectoral innovation on the instrument. Table A.14 demonstrates that the instruments have power in predicting sectoral innovation output both in the U.S. (column 1) and globally (column 2). In both models, we control for fixed effects at the cross-section and in the time series. From these models, we calculate sectoral innovation predicted by these tax credits, $\ln n_{it}^{TAX}$.

Table A.14. Predicting Sectoral Patent Count Using R&D Tax Credits

	United States	Global
$Y=$	ln(Patents)	ln(Patents)
	(1)	(2)
$ln(User\ Cost\ of\ R\&D\ Capital)$	-11.774***	-0.288**
	(4.041)	(0.134)
Fixed Effects		
Sector	Yes	
Year	Yes	
Country x Sector		Yes
Country x Year		Yes
Sector x Year		Yes
R^2	0.866	0.969
No. of Sectors	158	
No. of Country x Sectors		1,242
No. of Obs	4,615	18,799

Notes. This table presents evidence that the user cost of R&D capital predicts patent output. Standard errors are clustered at the sector and year levels.

In the main 2SLS analysis, for each sector, we calculate upstream knowledge using the same equation as in (26), replacing the realized sectoral innovation with the fitted values $\ln n_{it}^{TAX}$. We denote this fitted value of the knowledge as $\text{Knowledge}_{it}^{Up,TAX}$. The variable $\text{Knowledge}_{it}^{Up,TAX}$ is then used as an instrument in the analysis in (27). We report the first-stage regressions in Table A.15, and domestic and global versions of the knowledge diffusion results in Tables A.16 and A.17, corresponding to Tables 2 and 3 in the paper.

Table A.15. Predicting Sectoral Patent Count Using R&D Tax Credits

	United States	Global
$Y=$	$Knowledge_{it}^{Up}$ (1)	$Knowledge_{mit}^{Up}$ (2)
$Knowledge_{it}^{Up,IV}$	1.092***	
	(0.053)	
$Knowledge_{mit}^{Up,IV}$		0.542***
		(0.045)
$ln(R\&D\ Stock)_{i,t-1}$	0.060***	0.001
	((0.018))	(0.007)
Fixed Effects		
Sector	Yes	
Year	Yes	
Country x Sector		Yes
Country x Year		Yes
Sector x Year		Yes
F-statistics	427	148
R^2	0.984	0.982
No. of Sectors	95	
No. of Country x Sectors		282
No. of Obs	1140	4587

Notes. The first-stage regression, instrumental variable is the fitted value of upstream innovation accumulated through the innovation network. Standard errors are clustered at the sector and year levels.

Table A.16. US Evidence of Knowledge Spillovers Through Innovation Networks–*Second-Stage IV Results*

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{it}^{Up}$	0.583**	0.594**	0.591**	0.917***	0.931***	0.926***
	(0.269)	(0.269)	(0.265)	(0.289)	(0.288)	(0.287)
$ln(R\&D\ Stock)_{i,t-1}$	0.408***	0.424***	0.388***	0.206	0.225	0.183
	(0.111)	(0.123)	(0.107)	(0.133)	(0.146)	(0.134)
$Knowledge_{it}^{Down}$		-0.057			-0.068	
		(0.134)			(0.110)	
$Knowledge_{it}^{Up,IO}$			0.248			0.282
			(0.357)			(0.407)
R^2	0.169	0.169	0.171	0.092	0.093	0.090
No. of Sectors	95	95	95	95	95	95
No. of Obs	1140	1140	1140	1140	1140	1140
Fixed Effects		Sector, Year			Sector, Year	

Notes. Second-stage regression. Same setting as in Table 2.

Table A.17. Global Evidence of Knowledge Spillovers Through Innovation Networks–*Second-Stage IV Results*

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{mit}^{Up}$	0.226**	0.235**	0.255**	0.453***	0.462***	0.485***
	(0.113)	(0.106)	(0.116)	(0.143)	(0.142)	(0.149)
$ln(R\&D\ Stock)_{i,t-1}$	0.079***	0.079***	0.081***	0.083***	0.083***	0.085***
	(0.020)	(0.020)	(0.020)	(0.030)	(0.030)	(0.030)
$Knowledge_{mit}^{Down}$		-0.020			-0.024	
		(0.085)			(0.116)	
$Knowledge_{mit}^{Up,IO}$			-0.213			-0.205
			(0.390)			(0.392)
R^2	0.035	0.036	0.022	0.028	0.029	0.023
No. of Country x Sectors	282	282	277	282	282	277
No. of Obs	4587	4587	4527	4587	4587	4527
Fixed Effects	Country x Sector, Country x Year, Sector x Year					

Notes. Second-stage regression. Same setting as in Table 3.

E.4 Additional Results on R&D Misallocation

This subsection presents additional results that quantify R&D misallocation, supplementing Section 5.

- Tables A.18 and A.19 present cross-country and time-series correlations of optimal R&D allocation γ.
- Figure A.10 presents US model fits with some labeled sectors.
- Figure A.11 presents analysis using alternative parameters of ρ/λ; Figure A.12 presents analysis using data from different years.
- Figures A.13, A.14, and A.15 present analysis using patent outputs and OECD R&D expenditure shares as innovation allocation measures, supplementing analysis using R&D expenditure shares (aggregated from firm-level data) in the paper.
- Figure A.16 provides additional analysis on the time series of R&D misallocation and implied welfare cost.
- Table A.20 summarizes the robustness of our quantitative analysis across different specifications of Ω, ρ, and λ.
- Figure A.17 presents evidence on R&D misallocation within 1-digit IPC patent classes.

Table A.18. Unilaterally Optimal R&D Allocations Across Countries

Countries	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands	EU
US		0.97	0.90	0.93	0.95	0.84	0.94	0.94	0.92	0.95	0.95
Japan	0.91		0.93	0.94	0.96	0.87	0.94	0.94	0.93	0.94	0.96
China	0.87	0.93		0.95	0.91	0.91	0.91	0.90	0.91	0.90	0.94
South Korea	0.85	0.89	0.84		0.92	0.83	0.90	0.90	0.88	0.89	0.92
Germany	0.77	0.89	0.79	0.82		0.85	0.97	0.96	0.94	0.97	0.99
Russia	0.70	0.76	0.86	0.60	0.57		0.84	0.82	0.90	0.86	0.86
France	0.81	0.89	0.87	0.73	0.73	0.76		0.98	0.94	0.97	0.98
UK	0.84	0.89	0.86	0.73	0.73	0.76	0.99		0.94	0.97	0.98
Canada	0.78	0.88	0.88	0.72	0.71	0.84	0.97	0.96		0.95	0.95
Netherlands	0.83	0.89	0.87	0.74	0.72	0.76	0.98	0.97	0.96		0.97
EU	0.87	0.96	0.91	0.82	0.90	0.74	0.95	0.95	0.93	0.94	

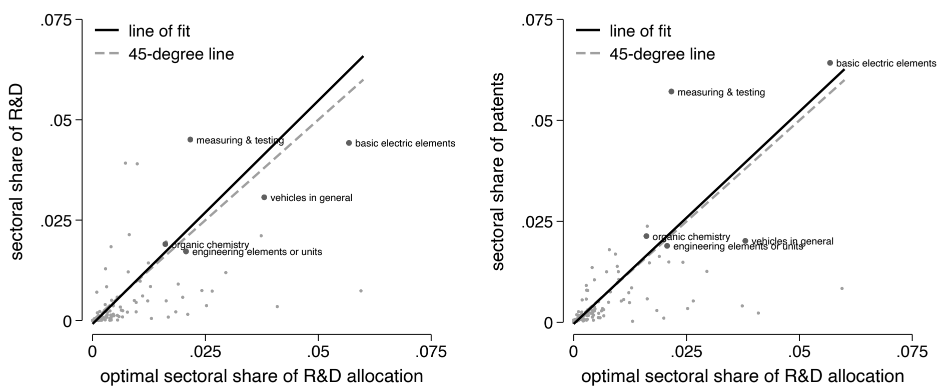
Notes. This table shows the pair-wise correlations of optimal R&D allocations γ across countries using country-specific statistics as of 2010. The lower triangular panel shows the Pearson correlation coefficients; the upper triangular panel shows Spearman's rank correlation.

Table A.19. Unilaterally Optimal US R&D Allocations of Across Time

Time Period	2020	2010	2000	1990	1980
2020		1.00	0.99	0.98	0.98
2010	0.99		0.99	0.98	0.98
2000	0.97	0.97		1.00	0.99
1990	0.96	0.94	0.99		1.00
1980	0.94	0.93	0.99	1.00	

Notes. This table shows the pair-wise correlations of optimal R&D allocations γ across different time periods using U.S. statistics during the specific year. The lower triangular panel shows the Pearson correlation coefficients; the upper triangular panel shows Spearman's rank correlation.

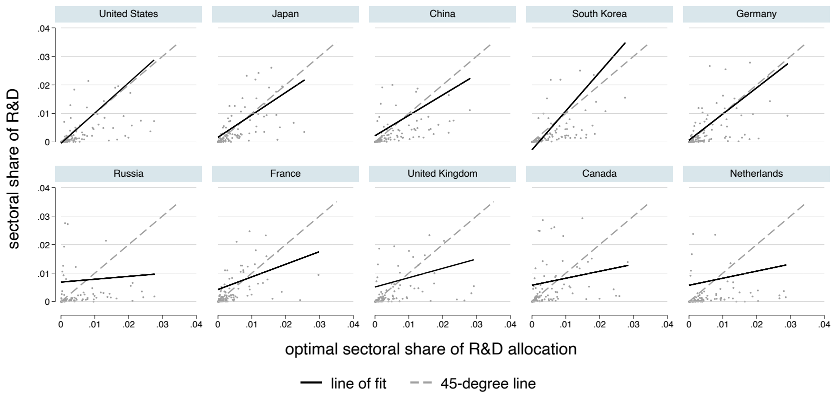
Figure A.10. U.S. Actual R&D Allocation vs. Optimal Allocation$\boldsymbol{\gamma}_{US}$



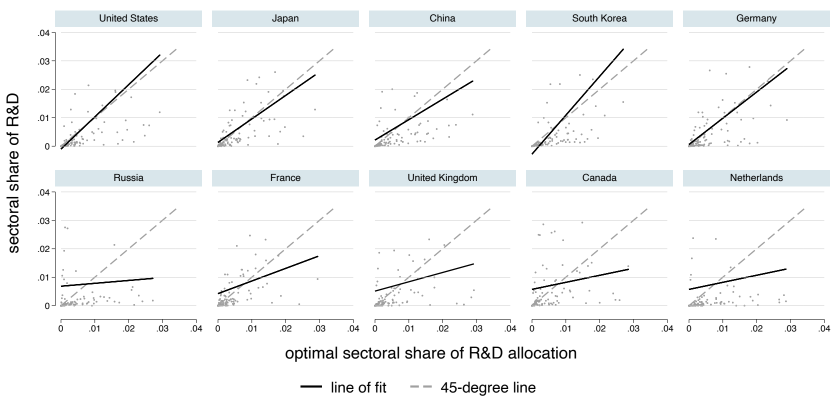
Notes. This figure adds sector labels to Figure 5 in the paper. It shows scatter plots of real-world sectoral R&D expenditure shares (left panel) and patent output shares (right panel) against optimal R&D allocation shares, γ_{US}, for the U.S. in 2010-2014. The solid line is the linear fit; the dashed line is the 45-degree line. For visual clarity, we exclude three outlier sectors that account for >7.5% of national R&D shares or national patent output from the scatter plots, but all sectors are used when constructing the linear fit.

Figure A.11. Alignment Between Real Allocation and Optimal Allocation Across Countries
Using Alternative Parameter Values

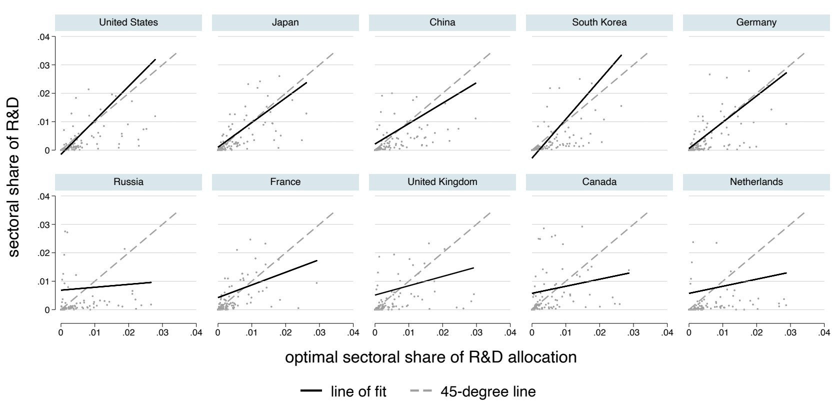
Panel (a): Use $(1+\rho/\lambda)^{-1} = 0.7$



Panel (b): Use $(1+\rho/\lambda)^{-1} = 0.8$



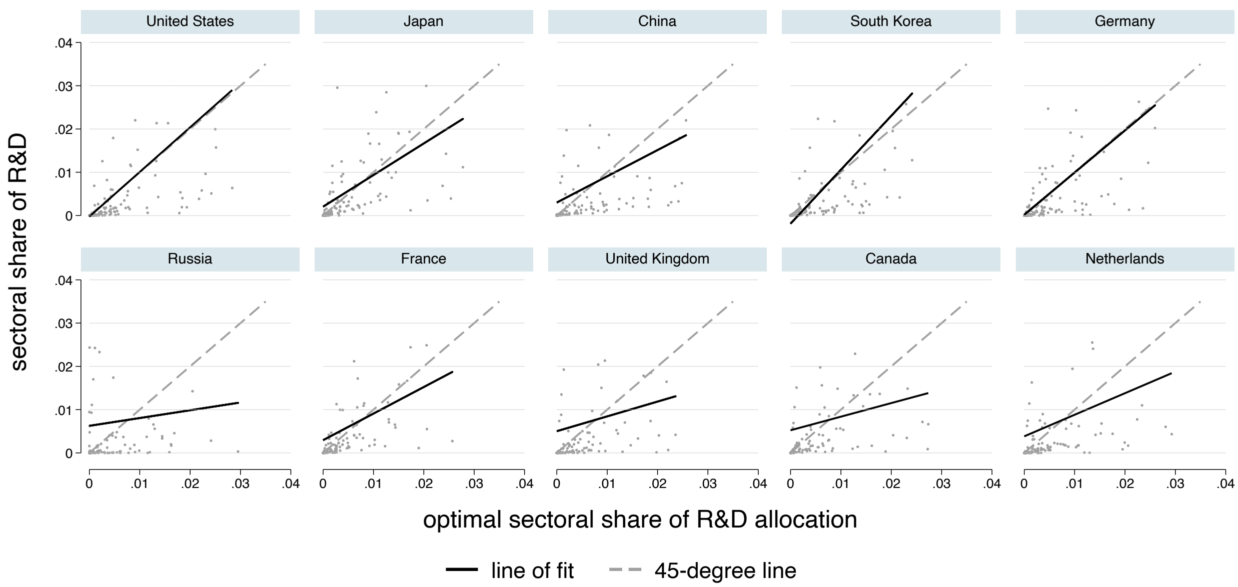
Panel (c): Use $(1+\rho/\lambda)^{-1} = 0.9$



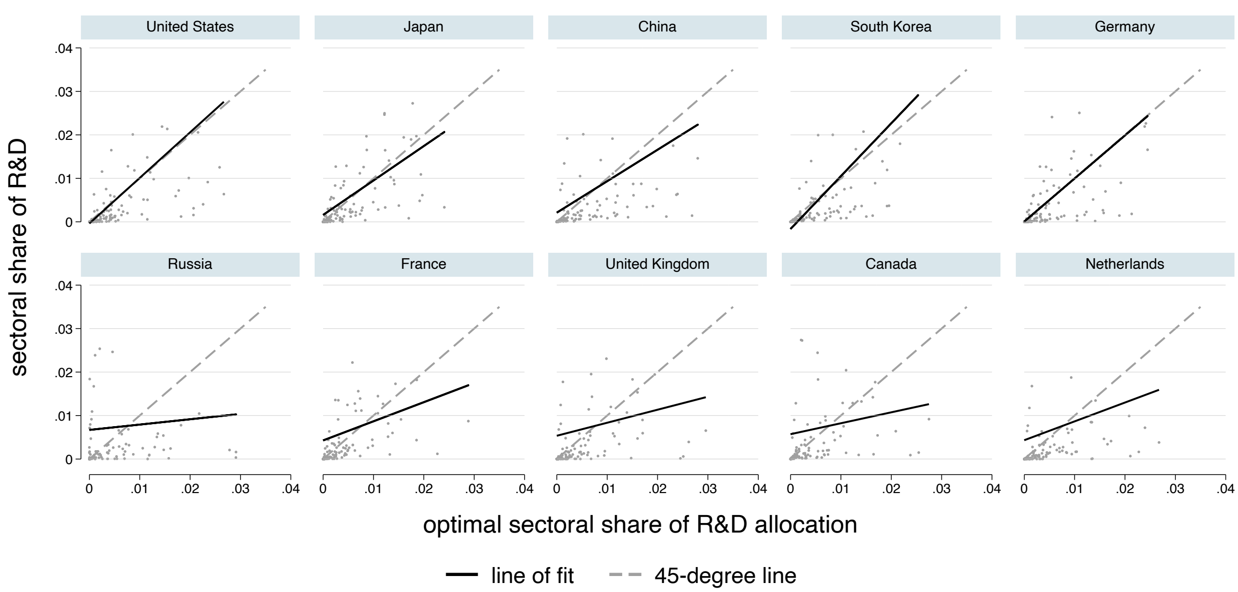
Notes. This table reproduces Figure 6 in the paper with alternative parameter values of ρ/λ.

Figure A.12. Alignment Between Real Allocation and Optimal Allocation Across Countries
Different Years

Panel (a): 2000

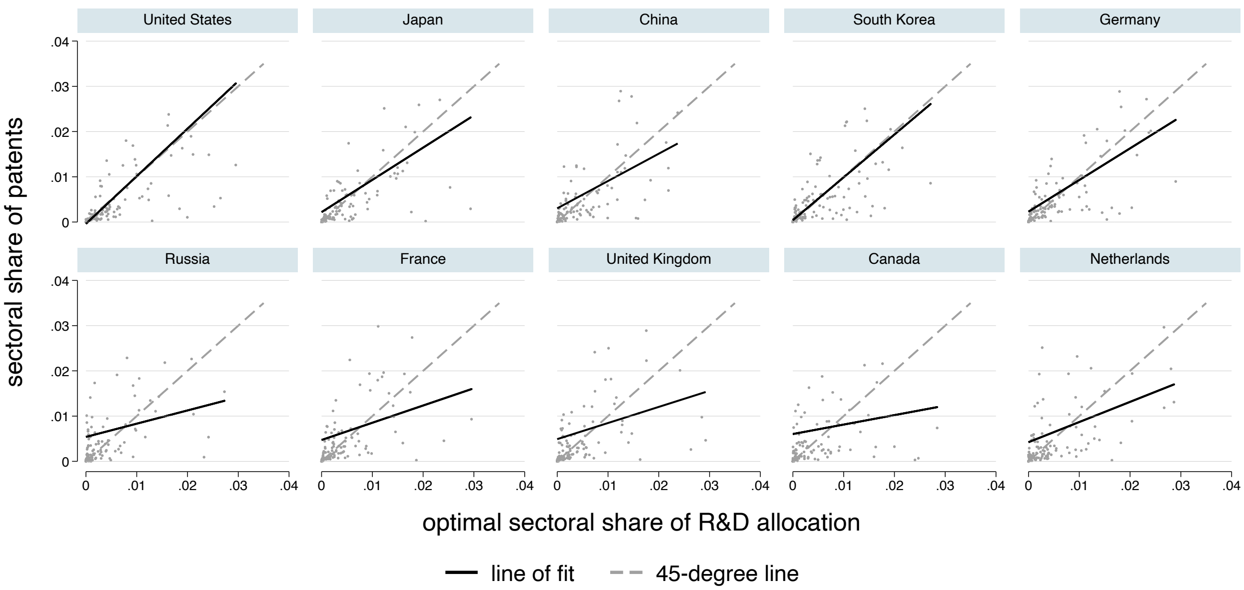


Panel (b): 2005



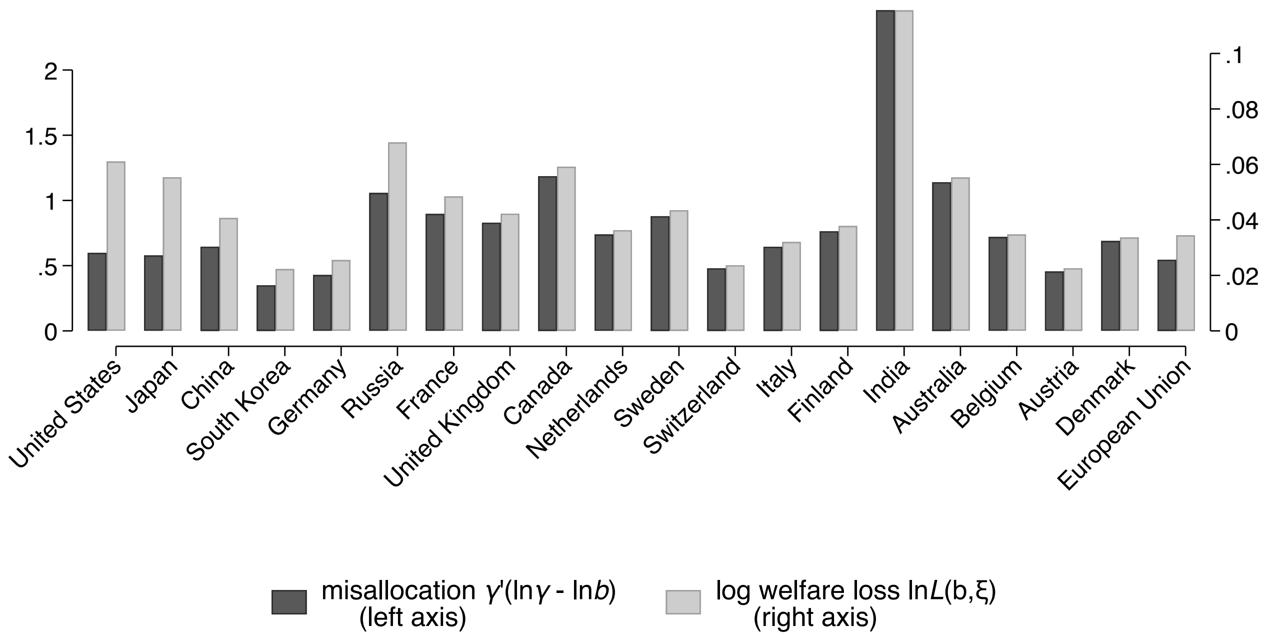
Notes. This figure reproduces Figure 6 in the paper using data from alternative years. The figure shows scatter plots of sectoral R&D expenditure share in total national R&D expenditures against the optimal sectoral share of R&D allocation for top ten innovative countries. The solid line is the linear fit; the dashed line is the 45-degree line.

Figure A.13. Alignment Between Real Allocation and Optimal Allocation Across Countries
Using Sectoral Share of Patents as Real Allocation



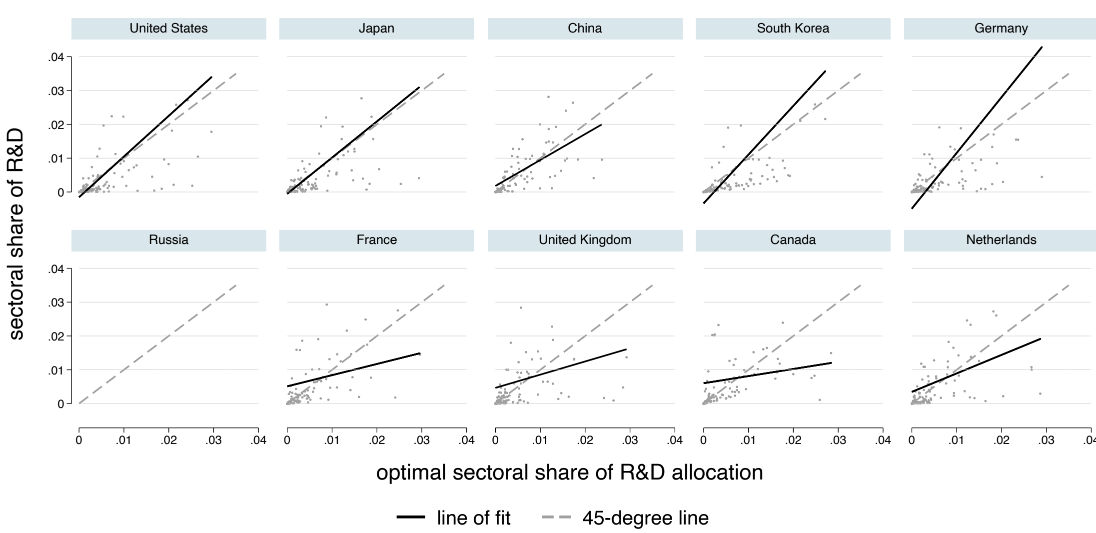
Notes. This figure reproduces Figure 6 in the paper. The figure shows scatter plots of sectoral patent output share in total patent output in the country against the optimal sectoral share of R&D allocation for top ten innovative countries in 2010. The solid line is the linear fit; the dashed line is the 45-degree line.

Figure A.14. Country-Level Welfare Loss from Misallocation
Using Sectoral Share of Patents as Real Allocation



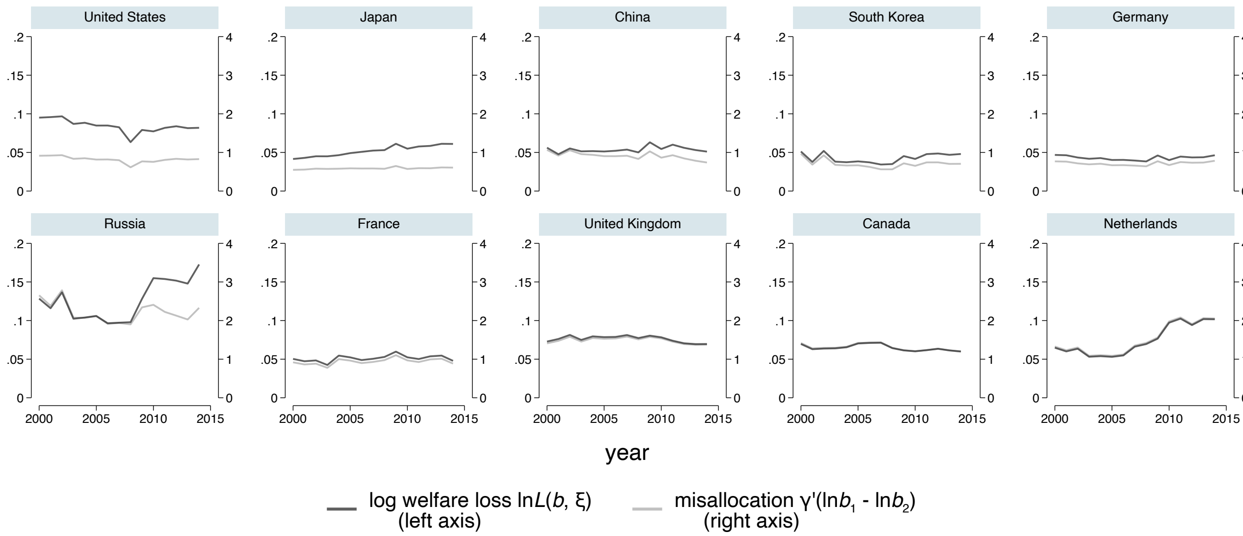
Notes. This table shows the level of R&D misallocation and associated welfare cost during 2010—2014. The table reproduces Figure 7 in the paper, but uses sectoral share of patents, rather than R&D expenditure shares, as the real allocation.

Figure A.15. Alignment Between Real Allocation and Optimal Allocation Across Countries
Using OECD R&D Shares as Real Allocation



Notes. This figure reproduces Figure 6 in the paper. The figure shows scatter plots of sectoral R&D share as reported in the OECD ANBERD database against the optimal sectoral share of R&D allocation for top ten innovative countries in 2010. The solid line is the linear fit; the dashed line is the 45-degree line.

Figure A.16. R&D Misallocation and Welfare Cost Across Countries and Over Time



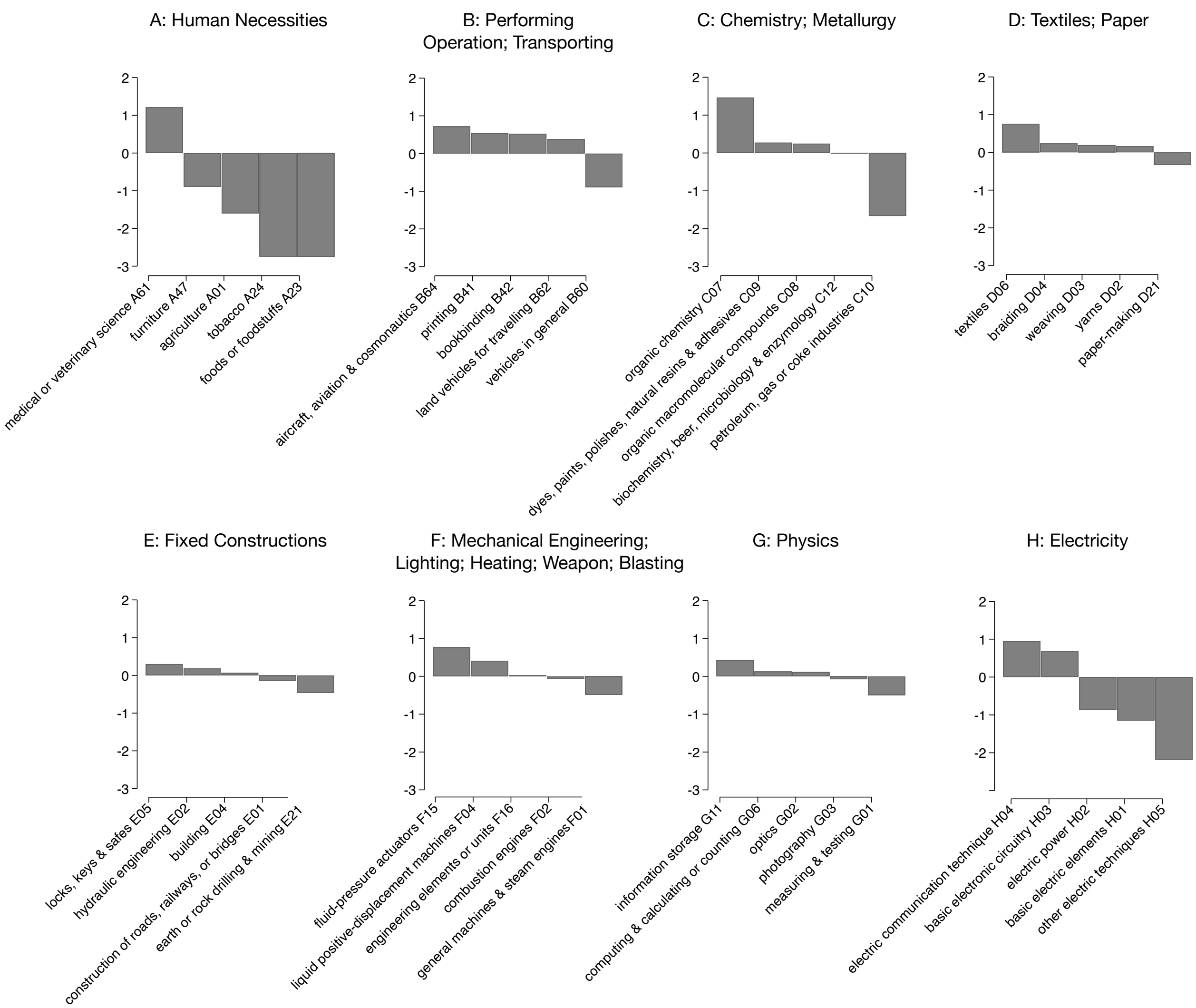
Notes. This figure plots the level of misallocation and welfare cost across countries over time. The calculation focuses on misallocation in top 50 IPC classes by total patents.

Table A.20. Robustness of γ and Centrality Across Specification of $\boldsymbol{\Omega}$, ρ, and λ

	Specifications	Average Correlation With Baseline Case			
		Optimal Allocation γ		Centrality a	
		Pearson's r	Spearman's ρ	Pearson's r	Spearman's ρ
A. Alternative Specifications of Ω					
A1	Forward-citation weighted Ω	0.9974	0.9994	0.8916	0.9864
A2	Backward-citation weighted Ω	0.9999	0.9999	0.9974	0.9968
A3	Scaled Ω	0.9955	0.9959	0.5228	0.9327
B. Alternative Values of ρ *and* λ					
B1	Using $(1+\rho/\lambda)^{-1} = 0.4$	0.9976	0.9984	-	-
B2	Using $(1+\rho/\lambda)^{-1} = 0.5$	0.9986	0.9990	-	-
B3	Using $(1+\rho/\lambda)^{-1} = 0.6$	0.9993	0.9996	-	-
B4	Using $(1+\rho/\lambda)^{-1} = 0.7$	0.9999	0.9998	-	-
B5	Using $(1+\rho/\lambda)^{-1} = 0.8$	1.0000	1.0000	-	-
B6	Using $(1+\rho/\lambda)^{-1} = 0.9$	0.9994	0.9997	-	-
B7	Using $(1+\rho/\lambda)^{-1} = 0.95$	0.9988	0.9994	-	-
C. Industry-Specific λ					
C1	ROA (median = 0.1747, s.d. = 0.0268)	0.9982	0.9994	-	-
C2	Gross Profit Margin (median = 0.2242, s.d. = 0.0460)	0.9985	0.9995	-	-
D. Injecting Measurement Errors into Ω					
D1	Adding log-N(0.02, 0.02) noise to Ω	0.9936	0.9934	0.8900	0.8166
D2	Adding log-N(0.04, 0.04) noise to Ω	0.9936	0.9934	0.8199	0.6607
D3	Adding N(0.02, 0.02) noise to Ω	0.9962	0.9944	0.9105	0.8192
D4	Adding N(0.04, 0.04) noise to Ω	0.9951	0.9931	0.8417	0.6696
D5	Adding $\max\{N(0.02, 0.02), 0\}$ noise to Ω	0.9963	0.9945	0.9118	0.8299
D6	Adding $\max\{N(0.04, 0.04), 0\}$ noise to Ω	0.9952	0.9932	0.8506	0.6941
D7	Adding U[0, 0.02] noise to Ω	0.9978	0.9967	0.9453	0.9287
D8	Adding U[0, 0.04] noise to Ω	0.9965	0.9953	0.9252	0.8861
D9	Adding Exp(0.02) noise to Ω	0.9964	0.9942	0.9068	0.8059
D10	Adding Exp(0.04) noise to Ω	0.9952	0.9928	0.8320	0.6597

Notes. This table reports the average Pearson and Spearman-rank correlation of γ and centrality of the innovation networks between the benchmark specification and different sets of alternative innovation network constructions for $\boldsymbol{\Omega}$ (Panel A), alternative values of ρ and λ (Panels B and C), and $\boldsymbol{\Omega}$ with injected errors (Panel D). In rows A1 and A2, we weigh each cite in $\boldsymbol{\Omega}$ construction (24) by the quality (total forward citations received) of either the citing or the cited patent. In row A3, we construct $\omega_{ij} \propto Cites_{i\to j}$ to scale directly with the total citations totally across or ij-pairs (rather than normalized by the citations from i), and we choose the proportionality constant so that the spectral radius of $\boldsymbol{\Omega}$ is equal to one, ensuring endogenous growth as in our baseline model. Rows B1 to B4 consider a range of alternative values for ρ and λ. Changing ρ/λ affects, across all countries, the magnitude of the welfare impact of R&D reallocation, but the cross-country welfare impact still correlates highly with our baseline specification. Panel C considers a specification with sector-specific innovation step size λ_i. Motivated by the decentralized economy constructed in Section 2.7.3, where the step size corresponds to the profit share, we measure λ_i using each sector's median ROA (return on assets) calculated from our firm-level datasets, and calculate the corresponding γ and welfare impact of R&D reallocation using the theoretical extension in Section B.8. Finally, in Panel D, we show our quantitative analysis is robust to introducing additional, simulated random errors to $\boldsymbol{\Omega}$. For each of the listed distribution, in each simulation, we add to each element in $\boldsymbol{\Omega}$ random and independent terms drawn from the distribution rescale $\boldsymbol{\Omega}$ to ensure row sum to be one. For each distribution, we simulate the exercise for 10,000 times, and the correlations are reported as average of the benchmark with each of the simulated $\boldsymbol{\Omega}$.

Figure A.17. U.S. R&D Misallocation within 1-digit IPC Classes



Notes. This figure shows U.S. R&D misallocation across 3-digit IPC classes within each 1-digit IPC class. Each of the eight 1-digit IPC categories is represented by a separate panel, in which we show the log-ratio between actual R&D and the constrained-optimal R&D allocation if a planner can reallocate resources across 3-digit IPC classes within the 1-digit IPC category.