Innovation Networks and R&D Allocation*

Ernest Liu† — *Princeton and NBER*

Song Ma‡ — *Yale and NBER*

August 16, 2022

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 path of R&D resource allocation. We show a planner valuing long-term growth should allocate more R&D toward upstream sectors in the innovation network, but the incentive is muted in open economies that rely on foreign knowledge spillovers. We derive sufficient statistics for evaluating the welfare cost of R&D misallocation. Empirically, we build the global innovation network based on patent citations and establish its empirical importance 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., reducing R&D misallocation down to Japan's level could generate more than 28% welfare gains.

Keywords: Innovation Network, Resource Allocation, Growth

JEL Classification: O33, O38, F43

*This paper subsumes the previous working paper entitled "Innovation Networks and Innovation Policy". We thank Daron Acemoglu, Philippe Aghion, Manuel Amador, Pol Antras, David Atkin, Ariel Burstein, Lorenzo Caliendo, Vasco Carvalho, Ben Golub, Ben Jones, Hugo Hopenhayn, Bill Kerr, Pete Klenow, Sam Kortum, Atif Mian, Ezra Oberfield, Alessandra Peter, Stephen Redding, Peter Schott, Kjetil Storesletten, Alireza Tahbaz-Salehi, Aleh Tsyvinski, John Van Reenen, Heidi Williams, and Kei-Mu Yi for useful feedback. We also thank seminar participants at Cambridge, EIEF, EUI, Minnesota, NBER ITI, NBER Summer Institute (Innovation), Princeton, UNC Chapel Hill, Yale, ZJU, EAGLS, College de France, SED, Global Economic Networks conference, Richmond Fed, and Philadelphia Fed. 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

Innovation is the source of long-run growth. How to foster innovation has long been a central question for economists and policy makers, and 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 R&D allocations differ across countries? What is the level of R&D misallocation across the world and how much gain does it create to improve 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 to assess the allocative efficiency of R&D in major economies around the globe. 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 important 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, quality-ladder 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. Despite the complexity of dynamic network spillovers, we are able to explicitly solve for the optimal path of cross-sector R&D resource allocation 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 for optimal allocation 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 upstream 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 innovation-downstream sectors such as consumer goods.

A key object is the innovation network's eigenvector centrality, which we call "*innovation centrality*." This captures the extent to which a sector's R&D activities contribute to economic growth, taking into account the network effects. We refer to sectors with higher innovation centrality as more fundamental in the innovation network. 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 consumer expenditure shares vector. 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 assess the level of R&D misallocation and calculate associated welfare costs. The consumption-equivalent welfare cost of cross-sector R&D misallocation, taking into account the economy's entire dynamic path, is 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 welfare cost of R&D misallocation. 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 evaluate the welfare cost of R&D misallocation.

We next extend our model to an open economy setting with international knowledge spillovers. We derive the *unilaterally* optimal R&D allocations that maximize each country's domestic welfare. The unilaterally optimal R&D allocations differ across countries not only because of differences in preferences, production structure, and the innovation network, but importantly, also because countries rely differentially on foreign knowledge spillovers. Holding the total level of R&D constant, an economy more reliant on foreign knowledge spillovers has less incentive to direct resources toward fundamental or innovation-upstream sectors and more incentive to direct resources toward consumer goods. Economies with well-developed domestic innovation networks, such as the U.S. and Japan, should conduct more R&D in innovation-central sectors; by contrast, economies that heavily rely on foreign knowledge spillovers should direct more R&D toward consumer goods.

The model provides sufficient statistics for the welfare cost of cross-sector R&D misallocation in open economies. Relative to the closed economy analogue, the open economy welfare cost also depends on the degree of foreign reliance: the more an economy relies on foreign spillovers, the less consequential domestic R&D misallocation is for domestic welfare.

Our empirical analysis starts by constructing a global innovation network from over 36 million patents and their citations, collected from over 40 main 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. Innovation centrality is highly heterogeneous across 131 3-digit international technological classes (IPCs). A handful of IPCs—such as medical science, computing, and semiconductors—should be allocated disproportionately large shares of R&D resources in order to maximize growth. 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, Germany, and, in early periods of the sample, Japan—are foreign-reliant, with domestic citation shares well below 50%.

We empirically validate the key mechanism behind our theory, namely that a sector's innovation activities benefit from past innovation in upstream sectors linked through the innovation network. We first show that a sector's patent production increases after its upstream sectors generate more patents. This pattern holds in both the U.S. domestic innovation network, as in Acemoglu, Akcigit, and Kerr (2016), and the global innovation network and across a variety of innovation output metrics, including patent counts, future citations, and patent value measured by stock market reactions upon patent approval. To ensure comovements in patent output are driven by knowledge spillovers and not by common shocks to connected sectors (Manski, 1993, Bloom, Schankerman, and Van Reenen, 2013), we verify our results using instrumental variables based on time-varying sectoral exposure to tax-induced user cost of R&D, exploiting changes in federal and state-specific rules. Additionally, knowledge spillovers are directional: 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. The innovation network only weakly correlates with the input-output production network, such that there is substantial independent variation in both network structures. Relative to input-output linkages, the innovation network is a significantly stronger channel through which knowledge spillovers take place.

Our main empirical application uses the model to evaluate cross-sector R&D allocations in the real world. For each country and year, 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. The model-implied optimal R&D allocation strongly predicts actual resource allocation: 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 welfare costs from R&D misallocation relative to each country's own efficient benchmark. We find Japan has the most efficient R&D allocation. For the period of 2010–2014, adopting the optimal allocation across the fifty largest IPCs (which jointly account for over 90% of patent output) could lead to welfare improvements equivalent to raising consumption by 26% at every point in time for Japan. Using Japan as a benchmark, we find that for 2010–2014, moving to Japan's level of R&D allocative efficiency can improve consumption-equivalent welfare by 28.8% in the U.S., 4.9% in China, 12.4% in South Korea, and 17.8% in Germany. These cross-country differences in R&D allocative efficiency are qualitatively stable over time except for China, which has improved R&D allocative efficiency since 2000. It is important to note that a more allocatively efficient economy is not necessarily more innovative in absolute terms; instead, our allocative efficiency measure reflects the distance of actual R&D allocation from each country's own efficient benchmark.

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. Our calculation shows that the technology class most relevant to semiconductor technologies (H01) is under-allocated in the U.S. (and while slightly over-allocated in China and South Korea), providing support to the recent U.S. initiatives to accelerate and catalyze domestic semiconductor sector. Under-allocated sectors also represent technology classes related to "green innovation" such as waste and pollution management and alternative energy. Though these exercises are suggestive, they show our model's potential to analyze and address more detailed R&D policy issues.

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, 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). 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. Our normative analysis distinguishes us from other works that consider cross-sector knowledge linkages, including Acemoglu et al. (2016), Cai and Li (2019), Huang and Zenou (2020), and Yang and Zhu (2020), and, in an open economy setting, Cai et al. (forthcoming) and Guillard et al. (2021). We provide simple sufficient statistics for the welfare cost of R&D misallocation, and we use these sufficient statistics to evaluate actual misallocation across countries. There is also a large literature on 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. Relative to this literature, which focuses on the country-level implications of foreign spillovers, the novelty of our open economy analysis is to show how sectoral-level foreign dependence interacts with the innovation network and shapes the unilaterally optimal R&D allocation across sectors.

Second, we contribute to the large literature on resource misallocation (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 such literature mostly focuses on the misallocation of production resources, we instead study the misallocation of innovation resources, and our analysis is therefore inherently dynamic in nature.

Third, we contribute to the fast-growing literature on general equilibrium networks (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., 2021b, vom Lehn and Winberry, 2022). Particularly related are recent papers that introduce methodologies for dynamic network analysis (Kleinman et al., 2021a, Liu and Tsyvinski, 2021) 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.

The remainder of the paper is structured as follows. Section 2 hosts our closed economy model, and Section 3 extends the analysis to an open economy setting. Section 4 introduces our data. Section 5 describes the global innovation network and provides evidence of its importance for knowledge spillovers. Section 6 hosts our main empirical application, where we use the model to evaluate cross-sector R&D allocations across countries and time. Section 7 concludes. A separate Online Appendix contains the derivations of the results in the paper, a number of further extensions, and supplementary materials on data and empirical results.

2 Theory: Innovation Network in A Closed Economy

We study the optimal allocation of R&D resources in a multisector, quality-ladder growth model with an innovation network. This section studies a closed economy. We set up the model in Section 2.1 and analyze the efficient allocation of R&D resources across sectors in Section 2.2. We analyze both optimal R&D allocation's long-run impact on the balanced growth path (Section 2.3) and the welfare impact taking into account the transitional dynamics (Section 2.4). In Section 2.5 we analyze a constrained problem in which the planner can reallocate resources only in a subset of sectors. In Section 2.6 we discuss potential inefficiencies in a stylized decentralized setting and how to implement the efficient allocation using R&D taxes and subsidies. Section 2.7 discusses other extensions in the Online Appendix.

2.1 Setup

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 produced by a Cobb-Douglas combination over sectoral composite 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 share of sector i.

Each sectoral composite good i is a Cobb-Douglas aggregator over a continuum of intermediate varieties. Each intermediate variety can be potentially supplied in a countably infinite number of qualities. Let $q_{it}(\nu)$ denote the highest quality of variety ν available in sector i. The sectoral composite good aggregator y_{it} is given by

$$\ln y_{it} = \int_0^1 \ln \left(q_{it}(\nu) \, x_{it}(\nu|q) \right) d\nu, \tag{3}$$

where $x_{it}(\nu|q)$ is the quantity of the variety ν of quality q in the production process. The sectoral aggregator (3) implicitly imposes that only the highest quality variety will be used in production.

The intermediate varieties are produced linearly, one-for-one from production workers:

$$x_{it}(\nu|q) = \ell_{it}(\nu) \text{ for all } i, t, \nu, q. \tag{4}$$

The Innovation Process R&D can improve product quality. Let q_{it} denote the average quality of the intermediate varieties used for production in sector i at time t:

$$\ln q_{it} \equiv \int_0^1 \ln q_{it}(\nu) \, \mathrm{d}\nu.$$

We refer to q_{it} as sector i's knowledge stock at time t. The collection of cross-sector knowledge stocks $\{q_{it}\}_{i=1}^{K}$ are the state variables of the economy.

At time t, mass s_i of scientists employed in sector i generate a *flow* of new innovation n_{it}:

$$n_{it} = s_{it}\eta_i\chi_{it}, \qquad \chi_{it} \equiv \prod_{j=1}^{K} q_{jt}^{\omega_{ij}}, \qquad \sum_{j=1}^{K}\omega_{ij} \equiv 1. \tag{5}$$

η_i is the exogenous component of innovation productivity, and χ_{it} is the endogenous component. Specifically, χ_{it} is a Cobb-Douglas combination of knowledge stock across all sectors. The aggregator χ_{it} 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 spread across all sectors of the economy. We impose the assumption that χ_{it} has constant returns to scale ($\sum_{j=1}^{K}\omega_{ij} = 1$) to ensure sustained and nonexplosive growth. Absent cross-sector knowledge spillovers, $\omega_{ij} = 1$ if $i = j$ and is zero otherwise.

New innovations stochastically translate into product quality improvements, thereby contributing to knowledge stocks. Specifically, we assume the innovation-induced quality change of each variety ν in sector i occurs following a Poisson process with arrival rate $\ln(n_{it}/q_{it})$. Upon arrival, a new vintage of the improved variety is discovered, with proportional quality improvement $\lambda > 0$. The new vintage thus has quality $(1+\lambda)\,q_{it}(\nu)$. Even though innovation-induced quality improvements are stochastic at the variety level, the law of motion for average quality is deterministic at the sector level. Sector i's knowledge stock evolves according to:

$$\dot{q}_{it}/q_{it} = \lambda \ln(n_{it}/q_{it}). \tag{6}$$

The arrival rate of quality improvements decreases in existing knowledge stock q_{it}, capturing the notion that breakthrough innovation is harder to find as the knowledge stock in sector i expands (Bloom et al., 2020). The key distinction between n_{it} and q_{it} is that the former is a *flow* variable reflecting R&D activities at time t, whereas the latter is a *stock* variable reflecting the accumulation of past innovations.

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 central 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 degree to which sector i's innovation production relies on sector j's existing knowledge. 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.

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 intermediate goods as in (4). Scientists are employed to conduct R&D. Let ℓ_{it} denote the total mass of workers employed in sector i; the market clearing conditions for production workers and scientists are:

$$\textstyle\sum_{i=1}^{K}\ell_{it} = \bar{\ell}, \qquad \ell_{it} \equiv \int_0^1 \ell_{it}(\nu)\, d\nu; \qquad \sum_{i=1}^{K} s_i = \bar{s}. \tag{7}$$

Remarks on Model Features

Remark 1. *Testable Implications of the Innovation Process.* The law of motion (6) for each sector's knowledge stock implies that the (log-) knowledge stock at a given time can be written as the discounted sum of (log-) past innovation flows:

$$\ln q_{it} = \lambda \textstyle\int_0^{\infty} e^{-\lambda s} \ln n_{i,t-s}\, \mathrm{d}s. \tag{8}$$

Taking logs of the innovation production function (5) and substituting for q_{it} using (8), we obtain a log-linear relationship between innovation flows in sector i, the amount of R&D resources employed in the sector, and past flows from upstream sectors in the network:

$$\ln n_{it} = \ln \eta_i + \ln s_{it} + \lambda \textstyle\sum_{j=1}^{K} \omega_{ij} \left(\int_0^{\infty} e^{-\lambda s} \ln n_{j,t-s}\, \mathrm{d}s \right). \tag{9}$$

Equation (9) is empirically testable. We measure new innovation flows n_{it} with patents. Equation (9) thus implies that, controlling for R&D expenditures, a sector tends to create more patents at times when its upstream sectors had more past patents, and the effect weakens over longer time lags. In Section 5.2 we test and show this relationship holds empirically.

Remark 2. *Sectors without Products.* In the model, sectors may represent technologies that enter preferences but lack marketable products (e.g., national defense), as well as technologies that do not directly enter preferences ($\beta_j = 0$) but have knowledge spillovers to others sectors (e.g., artificial intelligence technologies that facilitate R&D activities in other sectors).

Remark 3. *Input-Output (I-O) Linkages.* The baseline model features an innovation network $\boldsymbol{\Omega}$ in the form of cross-sector knowledge spillovers, without a production network of I-O linkages. In Section B.3 of the Online Appendix, we generalize our results to a setting that features both production and innovation networks, and we show, with straightforward modifications, our characterizations extend to that setting. Furthermore, we later empirically test equation (9) and show knowledge spillovers that occur through the innovation network dominate the potential spillovers through I-O linkages; for this reason, we abstract away from the production network in the baseline model.

Remark 4. *Separation of R&D and Production Resources.* In the baseline model, we specify that R&D and production require two distinct resource types: scientists $\bar{s}$ and production workers $\bar{\ell}$. We choose this specification for expositional simplicity; as we show below, our results characterize 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, our analysis of cross-sector allocation shares holds even in a richer model in which a single worker type can move between R&D and production (Section B.2 of the Online Appendix).

2.2 Efficient Allocation of R&D Resources

In this section we characterize the efficient 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

$$V^*\left(\{q_{i0}\}\right) \equiv \max_{\{\ell_{it}(\nu), s_{it}\}} \int_0^\infty e^{-\rho t} \sum_{i=1}^K \beta_i \ln y_{it} \, \mathrm{d}t, \tag{10}$$

subject to the sectoral aggregator (3) for y_{it}, production function of intermediate inputs (4), the flow of new innovation (5), law of motion for sectoral knowledge (6), and the resource constraints (7).

First, recognize from equations (3) and (4) 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 share vector $\boldsymbol{\beta}$: for all t, $\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell}$ 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 equations (3) and (4) to express consumption in terms of production worker allocation and then applying Lemma 1, we rewrite the planner's problem in vector form as

$$\max_{\{\gamma_t\}\ \text{s.t. } \gamma_t'\mathbf{1}=1\forall t} \int_0^\infty e^{-\rho t}\boldsymbol{\beta}' \ln \boldsymbol{q}_t \, \mathrm{d}t \tag{11}$$

$$\text{s.t. } \mathrm{d}\ln \boldsymbol{q}_t \big/ \mathrm{d}t = \lambda \cdot \left(\ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{\gamma}_t + (\boldsymbol{\Omega} - \boldsymbol{I}) \ln \boldsymbol{q}_t\right), \tag{12}$$

where we obtain (12) by substituting the innovation production function (5) into $\boldsymbol{q}_t$'s law of motion (6).

The planner's problem may seem intractable: the economy features an entire vector of state variables (sectoral knowledge stocks), and the law of motion involves dynamic network spillovers across sectors, meaning the allocation of scientists 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 (11) and the law of motion (12) 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.

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,*

$$\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}. \tag{13}$$

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 (13) 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. 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 innovation aggregator χ_{it} in equation (5) 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 the network spillovers occur through sectoral knowledge stock, the flow of new knowledge through current R&D activities can only affect innovative efficiency in the future. Hence, the importance of network effects in the optimal R&D allocation is modulated by the discount rate ρ relative to the innovation step size λ: 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 *effective 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 coincides with worker allocation. 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 are more downstream 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 allocations affect 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 ($\rho/\lambda \to 0$) and converges to the consumption share vector $\boldsymbol{\beta}$ when the planner is completely myopic ($\rho/\lambda \to \infty$).

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

$$\boldsymbol{a}' = \boldsymbol{a}'\boldsymbol{\Omega} \quad \text{and} \sum_{i=1}^{K} a_i = 1.$$

Because $\boldsymbol{\Omega}$ is a row-stochastic matrix, the innovation centrality vector $\boldsymbol{a}$ exists and is generically 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 $\sum_{i=1}^{K} b_i = 1$.

[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.

Lemma 2. *Consider a BGP in which the aggregate consumption and knowledge stock in every sector grow at the same 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, the economic growth rate along the BGP is*

$$g(\boldsymbol{b}) = \mathit{const} + \lambda \cdot \boldsymbol{a}' \ln \boldsymbol{b}, \tag{14}$$

where the constant on the right-hand side is equal to $\lambda \cdot (\ln \bar{s} + \boldsymbol{a}' \ln \boldsymbol{\eta})$.

Lemma 2 analytically expresses the growth rate of knowledge stock along a BGP as a function of the R&D allocation, $\boldsymbol{b}$. The endogenous component of the growth rate is the step size of innovation (λ) times the inner product between the innovation centrality ($\boldsymbol{a}$) and the vector of log-R&D allocation shares ($\ln \boldsymbol{b}$). The exogenous component (the constant term) on the right-hand side of (14) shows the growth rate is higher when scientists are more abundant (higher $\bar{s}$) or generate new innovations more productively (higher $\boldsymbol{\eta}$) and when innovations lead to quality improvements with a bigger step size (higher λ).

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

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 R&D allocations to maximize the economy's growth rate. 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 $\boldsymbol{\gamma}' = \frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$ as the solution to the following fixed point equation, which usefully shows how $\boldsymbol{\gamma}'$ varies with the effective discount rate ρ/λ:

$$\boldsymbol{\gamma}'(\boldsymbol{I} - \boldsymbol{\Omega}) + \frac{\rho}{\lambda}(\boldsymbol{\gamma}' - \boldsymbol{\beta}') = \boldsymbol{0}'. \tag{15}$$

The two terms on the left-hand side represent innovation centrality $\boldsymbol{a}$ and consumer preferences $\boldsymbol{\beta}$ as two determinants of the efficient R&D allocation vector $\boldsymbol{\gamma}$. When the first term is equal to the zero vector ($\boldsymbol{\gamma}'(\boldsymbol{I} - \boldsymbol{\Omega}) = \boldsymbol{0}'$), it must be the case that $\boldsymbol{\gamma}$ coincides with the innovation centrality $\boldsymbol{a}$. When the second term is the zero vector ($\frac{\rho}{\lambda}(\boldsymbol{\gamma}' - \boldsymbol{\beta}') = \boldsymbol{0}'$), it must be the case that $\boldsymbol{\gamma} = \boldsymbol{\beta}$. Under the optimal R&D allocation, the sum of the two terms must be equal to the zero vector, and ρ/λ modulates the relative importance of these two terms. When ρ/λ is close to zero—a patient social planner—the first term dominates ($\lim_{\rho/\lambda \to 0} \boldsymbol{\gamma} = \boldsymbol{a}$). When ρ/λ is large—an impatient planner—consumer preferences dominate ($\lim_{\rho/\lambda \to \infty} \boldsymbol{\gamma} = \boldsymbol{\beta}$).

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 share vector:* $\lim_{\rho/\lambda \to \infty} \boldsymbol{\gamma} = \boldsymbol{\beta}$.

2.4 Welfare Cost of R&D Misallocation

We now characterize the welfare cost associated with R&D misallocation, taking into account the dynamic effects of R&D along the economy's entire future path.

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

$$V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \widetilde{\boldsymbol{b}}\right) - V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \boldsymbol{b}\right) = \frac{\lambda}{\rho^2} \boldsymbol{\gamma}' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right). \tag{16}$$

Proposition 3 shows that the welfare differences 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-differences in R&D allocation vectors, multiplied by the scalar λ/ρ^2. The result holds for any time path of worker allocations 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 Cost of R&D Misallocation) 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 cost of R&D misallocation*, 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, while holding worker allocation unchanged.

The scalar $\mathcal{L}(\boldsymbol{b})$ quantifies the welfare gains that could be achieved by reallocating R&D resources optimally. Note that, because the flow output is log-additive in the knowledge stock $\boldsymbol{q}_t$ and worker allocation $\boldsymbol{\ell}_t$ (equations 3 and 4), $\mathcal{L}(\boldsymbol{b})$ does not depend on the path of worker allocations. We next provide a simple formula for $\mathcal{L}(\boldsymbol{b})$.

Proposition 4. *The consumption-equivalent welfare cost of R&D misallocation $\boldsymbol{b}$ is*

$$\mathcal{L}(\boldsymbol{b}) = \exp\left(\frac{\lambda}{\rho} \boldsymbol{\gamma}' \left(\ln \boldsymbol{\gamma} - \ln \boldsymbol{b}\right)\right). \tag{17}$$

Proposition 4 shows that the consumption-equivalent welfare cost of R&D misallocation $\boldsymbol{b}$ is, 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 λ/ρ, the inverse of the effective discount rate. The inner product term is a statistical distance measure of how $\boldsymbol{b}$ differs from $\boldsymbol{\gamma}$. Also known as the relative entropy (or the Kullback-Leiber divergence), it is equal to the expected log-differences between $\boldsymbol{\gamma}$ and $\boldsymbol{b}$, with the expectation taken using the efficient allocation as the distribution. Note that ρ/λ affects welfare through not only the proportionality constant but also the optimal allocation $\boldsymbol{\gamma}$ (as in Proposition 1): a more patient planner allocates more resources toward more innovation-upstream sectors with higher innovation centrality.

2.5 Constrained Optimal 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 sufficient statistics for the welfare cost of R&D misallocation 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 5. *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 (10). 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 cost of following R&D allocation $\boldsymbol{b}$ instead of the constrained-optimal allocation is $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\lambda}{\rho} \left(\sum_{j \in \mathcal{K}} \gamma_j \right) \left(\boldsymbol{\gamma}^{\mathcal{K}} \right)' \left(\ln \boldsymbol{\gamma}^{\mathcal{K}} - \ln \boldsymbol{b}^{\mathcal{K}} \right) \right)$.*

Proposition 5 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 consumption-equivalent welfare cost of R&D misallocation, 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 extent of misallocation among sectors in $\mathcal{K}$. Relative to the welfare cost formula (17) 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.

2.6 Illustrating the 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 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 $\boldsymbol{\beta}$ along a BGP, which coincides with the planner's solution only if the society is completely myopic ($\rho/\lambda \to 0$).

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. A main advantage of our results from previous sections is that we can measure optimal R&D allocations from the data, and, by comparing optimal and actual R&D allocations, we can assess welfare consequences of R&D misallocation without taking a stance on the market structure that leads to real-world R&D allocations.

Consider a decentralized economy in which each intermediate variety is produced by a distinct monopolist. Different vintages of the same variety are perfect substitutes. Because the most recent vintage's quality is λ proportionally higher than the next best vintage, the monopolist conducts limit pricing and charges a markup $(1 + \lambda)$. No vintages with dominated quality are produced in equilibrium.

In each sector, innovation is carried out by a continuum of potential entrants, who hire scientists to conduct R&D and generate new innovations according to equation (5). New innovation flows improve the quality of a random variety within sector i at Poisson rate $\ln(n_{it}/q_{it})$; the innovating firm overtakes as the monopolist of that variety until another successful innovation occurs in the future.

The representative consumer receives all workers' and scientists' income and firm profits. 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.

Proposition 6. *In the decentralized BGP, the allocations of R&D and production resources both follow the consumption shares:* $\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell}$ *and* $s_{it} = \beta_i \bar{s}$.

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 (5) thus implies that R&D allocation must follow $s_{it} = \beta_i \bar{s}$ along the BGP.

The decentralized R&D allocation $\boldsymbol{\beta}$ stands in contrast to the socially optimal allocation, $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} + \frac{\boldsymbol{\Omega}}{1+\rho/\lambda} + \left(\frac{\boldsymbol{\Omega}}{1+\rho/\lambda} \right)^2 + \cdots \right)$. 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.

Policy Implementation of the Optimal R&D Allocation The difference between the socially optimal and decentralized R&D allocations arises because decentralized R&D allocation is driven by pursuit of profits, as firms do not fully internalize subsequent knowledge spillovers from their own innovative activities. Given a broad set of tax instruments, the planner may have many equivalent ways to implement the optimal R&D allocation. The most direct implementation is through sector-specific R&D subsidies and taxes. Intuitively, the planner should tax R&D activities in sectors with high β_i (which encodes market incentives) relative to γ_i (which encodes social incentives) and subsidize R&D activities in sectors with low β_i relative to γ_i. Formally, suppose the planner has access to R&D tax instruments such that firms pay $(1 + t_i)$ times the wage rate of each scientist in sector i. Suppose the planner has access to lump-sum taxes on the consumer in order to balance its budget.

Proposition 7. *A decentralized BGP allocates R&D resources efficiently under sectoral R&D subsidies $1 + t_i \propto \frac{\gamma_i}{\beta_i}$, when the planner uses the appropriate lump-sum tax levied on the consumer to balance the budget.*

2.7 Other Closed Economy Extensions

The Online Appendix includes a number of additional extensions to our closed economy model. In Section B.2, we allow for factor mobility between production and R&D, and we show the optimal share of resources allocated to each sector continues to be characterized by results from Section 2.2. In Section B.3, we incorporate a production network into the model and show the term $\boldsymbol{\beta}'$ in the formula for optimal R&D allocation should be replaced by sectoral value-added as a share of GDP; otherwise, our characterization continues to hold. In Section B.4.1 we relax the Cobb-Douglas assumption on both goods and innovation production functions and show our key result on the optimal R&D allocation $\boldsymbol{\gamma}$ still holds along a balanced growth path.

3 Theory: Innovation Network in An Open Economy

In this section, we generalize our closed economy model to an open economy receiving knowledge spillovers from abroad. We analyze the problem of a planner choosing allocations to maximize welfare for its citizens, taking the time path of foreign knowledge stocks and terms of trade as given. Analogous to our closed economy analysis, we derive the unilaterally optimal R&D allocations that can be measured from data on production and cross-country patent citations, and we derive the formula for the open economy welfare cost of R&D misallocation. Compared to the closed economy model, the key new insight here is that in open economies benefiting from foreign knowledge spillovers, planners have muted incentives to invest in innovation-upstream technologies. Reliance on foreign knowledge spillovers is therefore a key determinant of the unilaterally optimal R&D allocation.

3.1 Setup

The open economy model extends the closed economy model in Section 2.1 by introducing international knowledge spillovers and trade. We continue to assume a representative consumer with log flow utility and exponential discounting. 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 are a Cobb-Douglas aggregator over sectoral composite goods, which are aggregations of intermediate varieties produced from labor (equations 2, 3, and 4). 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 denote the relative prices of foreign goods. Trade balance implies:

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

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

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

Like in the closed economy counterpart (5), n_{it} is the flow of new innovation generated by scientists s_{it} employed in sector i at time t; new innovation leads knowledge accumulation according to (6). χ_{it} is the endogenous component of R&D efficiency in sector i. In the open economy, 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 (5).

The planner's problem is allocating workers and R&D resources to maximize domestic welfare while 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), (3), and (4); the law of motion for domestic knowledge (6); and resource constraints (7). 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.

3.2 Optimal R&D Allocation and Welfare in An Open Economy

We now characterize the planner's problem (20). In Section 3.3 below we discuss various theoretical extensions to the open economy model.

Proposition 8. *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; $\circ$ represents the Hadamard product (i.e., element-wise multiplication).

Proposition 8 generalizes the results 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 (c.f. Proposition 1), each round of network effect is no longer captured by the innovation network $\boldsymbol{\Omega}$ but is instead captured by $\boldsymbol{\Omega}\circ\boldsymbol{X}$. This is because in the presence of knowledge spillovers from abroad, the domestic planner recognizes that domestic R&D only partially contributes 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., $\frac{\partial \ln \chi_i}{\partial \ln q_j} = \omega_{ij}x_{ij}$). In the data, we measure x_{ij} by the share of citations from sector i to sector j that are toward prior domestic patents. Proposition 8 thus allows us to assess R&D allocation's optimality in an open economy.

The proportionality constant ξ, which ensures γ_i sums to one, is a measure of R&D self-sufficiency. ξ plays an important role in our welfare analysis below. $\xi \leq 1$ in open economies and is increasing in x_{ij}; $\xi = 1$ only if the economy does not benefit from any foreign spillovers ($x_{ij} = 1$ for all i, j). When the domestic contribution to knowledge spillovers is constant at $x_{ij} = x$ across sector-pairs, $\xi = \frac{\rho}{\rho+\lambda(1-x)}$.

Cross-Country Implications Proposition 8 highlights that countries with more self-contained innovation networks should allocate more R&D to innovation-central sectors that generate more knowledge spillovers to other sectors. Conversely, countries with higher reliance on foreign knowledge spillovers should direct R&D toward sectors that account for greater domestic value-added. In other words, 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; that is, when a country's future innovations benefit more heavily from foreign spillovers, the planner allocates R&D resources as if they place less value on long-term innovation spillovers. The intuition continues to hold when x_{ij} is heterogeneous across sector-pairs: economies with 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 optimally allocate more R&D resources to sectors that create more network externalities; by contrast, countries that rely on foreign knowledge spillovers should optimally allocate R&D resources more myopically, focusing disproportionately on sectors with higher value-added (β_i).

Welfare Cost of R&D Misallocation in An Open Economy Our next result provides the welfare cost of R&D misallocation in an open economy, extending our closed economy result in Proposition 4. We again express the welfare cost in consumption-equivalent terms.

Proposition 9. *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 cost of R&D misallocation is*

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

Relative to the closed economy welfare cost formula (17), the open economy formula (22) has the additional term ξ, the R&D self-sufficiency measure defined in Proposition 8. Intuitively, domestic R&D affects welfare both directly by improving product quality and indirectly via knowledge spillovers. The more an economy relies on foreign knowledge spillovers (lower x_{ij}s and thus lower ξ), the less important domestic spillovers are, and thus domestic R&D misallocation becomes less consequential for consumer welfare. Note that $\left\{\boldsymbol{q}_t^f\right\}$ and $\{\boldsymbol{\ell}_t\}$ do not appear in the welfare cost formula (22); as we show in the proof in the Online Appendix, this is because consumer welfare is log-separable in domestic R&D, foreign knowledge, and worker allocations. Moreover, for any import prices, a proportional expansion in domestic output can translate one-for-one into additional consumption of both domestic and foreign goods; hence $\left\{p_t^f\right\}$ does not appear in (22) either. We later apply Proposition 9 to the data and compute the welfare cost of R&D misallocation across countries.

3.3 Open Economy Extensions

In Section B.4.2 of the Online Appendix, we show our key result regarding the optimal R&D allocation $\boldsymbol{\gamma}$ still holds along a balanced growth path after relaxing the Cobb-Douglas assumption on both the goods and the innovation production functions. In Section B.5, we show our characterization in Section 2.5, for settings in which the closed economy planner can only allocate R&D in a subset $\mathcal{K}$ of sectors, applies to the open economy model as well. The constrained optimal R&D allocation should be proportional to the corresponding entries in the unconstrained optimal allocation, and—as in the closed economy setting—the consumption-equivalent welfare cost has an additional term $\sum_{j\in\mathcal{K}}\gamma_j \leq 1$ reflecting the fact that there is less to be gained when the planner can reallocate resources in fewer sectors. In Section B.6 of the Online Appendix, we further extend the analysis to an open economy planner who considers domestic R&D's effect on foreign variables.

4 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, consumption, trade, and R&D. Here we briefly describe how we construct and harmonize these data. Section C of the Online Appendix provides more details.

4.1 Innovation Data

U.S. Patents U.S. patent data are from the United States Patent and Trademark Office (USPTO).[2] 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

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

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.

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 main 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 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 equal weights.

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.

4.2 Data on Production and R&D

In our cross-country analysis, for each country and sector, we use the World Input-Output Database (WIOD, Timmer et al., 2015) to obtain data on value-added, employment, revenue, intermediate inputs, value used for consumption, imports, and exports. 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. 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).

We also use three widely used firm-level data sets, Compustat, Worldscope, and Datastream, to obtain additional information such as R&D expenditure that can be aggregated to the country-sector-level. 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.

4.3 Concordances and Measurements

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 allows us to construct sectoral measures of innovation activities and reversely to project sectoral measures into technology class levels. In our empirical analysis, we use the sample before 2014 to mitigate the right-truncation problem, since patents filed more recently may still be in the approval process.

To map the innovation measure to sectors, we leverage publicly traded firms' sectoral classifications. 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.[3] 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. 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 measure innovation input, we aggregate R&D expenditures to the country-sector-year level from firm-level information in Compustat, Worldscope, and Datastream. To capture actual patent timing, we use the year a patent was filed rather than granted.

Some of our subsequent analyses are at the level of IPC classes. To reverse-project country-sector-year level measures (most notably value-added and R&D expenditures) 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 public firm innovation measures in Section C of our Online Appendix.

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

5 Innovation Network and Knowledge Spillovers

In this section, we first construct the innovation network $\mathbf{\Omega}$ and discuss its empirical properties. We then empirically establish a key mechanism in our model, that knowledge spillovers occur through innovation networks both domestically in the U.S. and globally.

5.1 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. We define ω_{ijt} as the share of total citations patents in sector i made to sector j in year t:

$$\omega_{ijt} \equiv \frac{Cites_{ijt}}{\sum_{k=1}^{K} Cites_{ikt}}. \tag{23}$$

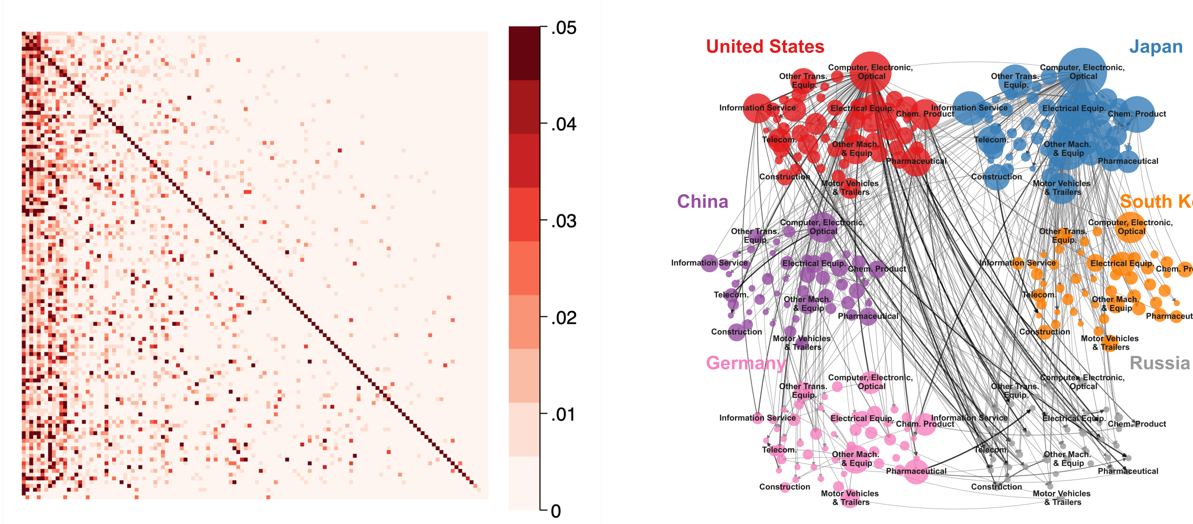
The object ω_{ijt} measures the extent to which upstream sector j's prior knowledge benefits innovation in sector i. The matrix $\mathbf{\Omega}_t$, whose ij-th entry is ω_{ijt}, captures the knowledge flow network we refer to as the *innovation network*. In a global setting, all subscript will include additional country dimensions indicating the countries of section 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 major innovation countries, suggesting their ability to capture some deep relations between sectoral innovation and technology diffusion. Table A.4 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.5 of the Online Appendix demonstrates that the innovation network constructed by pooling patents from all countries almost 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. Such high correlations mean that decisions about country and time specificities of the innovation network do not affect our results. Therefore, for expositional simplicity, we adopt the time-specific, location-invariant measure in (23) as our baseline notion of the innovation network.

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

Figure 1. Visualizing the Innovation Network

(a) IPC-to-IPC (131×131) network $\mathbf{\Omega}$ **(b)** The global innovation network



Notes. The left panel visualizes the IPC-to-IPC (3-digit level) network $\mathbf{\Omega}$ 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.

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.

The Innovation Network Weakly Correlates with Input-Output Networks The innovation network $\boldsymbol{\Omega}$ 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 $\boldsymbol{\Omega}$) commonly used to represent input-output relationships. Table 1 presents the correlation between entries in $\boldsymbol{\Omega}$ 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

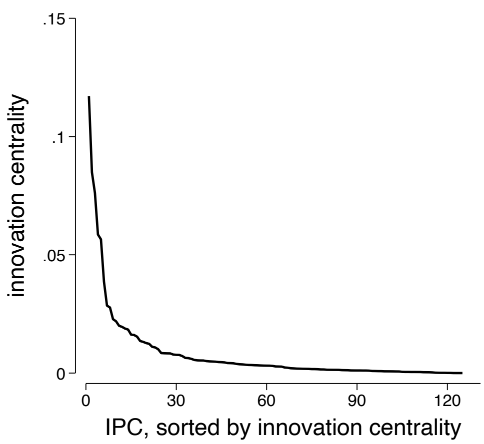
Notes. This table presents the correlations between the country-level innovation network matrix and the country-level input-output expenditure share matrix for the top 10 countries ranked by total patent counts during 2010–2014.

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 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 nine 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.

Cross-Country Linkages in the Innovation Network To the extent that the innovation network captures knowledge spillovers, how much do countries benefit from foreign knowledge? To answer this, for each country m, sector i, and year t, we compute the share of all citations mit made to patents in country m. 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

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

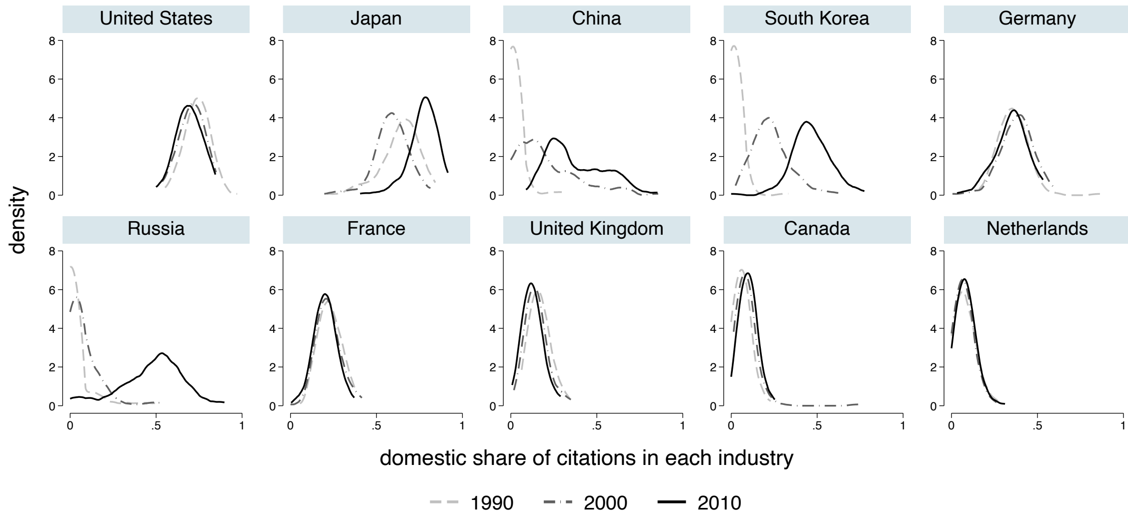
Notes. This figure presents the innovation centrality of different technology classes categorized using IPC. Panel (a) plots innovation centrality a_i across 3-digit IPC sectors ranked in descending order based on a_i. Panel (b) lists the top 10 IPC classes by their innovation centrality.

these three decades, about 70% of citations by U.S. patents are 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, the two other fast-growing Asian economies shown in Figure 3, although their levels of foreign reliance remain high.

5.2 Knowledge Spillovers Through the Innovation Network

We now provide evidence that knowledge spillovers occur through the innovation network. This evidence serves as a validation of both our theory's key mechanism and our measurement of the innovation network using patent citations. As we show, the knowledge spillover pattern holds across a variety of innovation output metrics in both the U.S. domestic network (as in Acemoglu et al., 2016) and the global innovation network. We also demonstrate that the spillovers are directional in nature—from upstream sectors to downstream, and not the reverse. Importantly, our findings are robust to using instrumental variables constructed based on time-varying sectoral exposure to tax-induced user cost of R&D to shock upstream innovation. We establish that the innovation network only weakly correlates with the input-output production network; the innovation network is a significantly stronger channel through which knowledge spillovers take place. Section E.2 in the Online Appendix presents more supplementary results.

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



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

5.2.1 U.S. Evidence

We first test the mechanism in the U.S., treating it as a closed economy. Specifically, our law of motion (6) implies that the logged knowledge stock in each sector is a discounted sum of logged patents as in (8), and the innovation production function (5) further implies a log-linear relationship (equation 9, reproduced below) between sector i's new patents, sectoral R&D, and past patents from other upstream sectors:

$$\ln n_{it} = \ln \eta_i + \ln s_{it} + \underbrace{\lambda \sum_{j=1}^{K} \omega_{ij} \left(\int_0^\infty e^{-\lambda s} \ln n_{j,t-s} \, \mathrm{d}s \right)}_{\substack{\equiv \ln \chi_{it}\text{, the knowledge stock that benefits} \\ \text{idea generation in sector } i \text{ at time } t}} . \tag{9}$$

This indicates 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}, and the effect decays over longer time lags with rate λ. Equation (9) 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 (9) by constructing the knowledge aggregator χ_{it} from past patents. 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 ($1 \leq \tau \leq 10$),[4] weighted by $\omega_{ij,t-\tau}$, the strength of the knowledge connection from j to i in the corresponding year:

$$\text{Knowledge}_{it}^{Up} \equiv \sum_{j \neq i} \sum_{\tau=1}^{10} \omega_{ij,t-\tau} \ln n_{jt-\tau}. \tag{24}$$

$\text{Knowledge}_{it}^{Up}$ captures the stock of past knowledge "upstream" of sector i, meaning it is the stock of knowledge ($\ln \chi_{it}$) that can benefit sector i's subsequent innovation production.

We then perform the following regression:

$$\ln n_{it} = \beta_1 \times \text{Knowledge}_{it}^{Up} + \beta_2 \times \ln R\&D_{i,t-1} + \xi_i + \xi_t + \text{controls}_{it} + \epsilon_{it}, \tag{25}$$

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 expenditure, with a one-year lag to reflect the delayed nature of patent filing (our results are robust to controlling for concurrent or additional lags of R&D expenditures). To purge time-invariant sectoral factors that affect patent output, we control for sector fixed effects ξ_i, and to purge time-varying shocks common across all sectors, we also control for year fixed effects ξ_t.

Before showing results, we will discuss two details. First, when constructing the upstream knowledge aggregator (24) for each sector i, we exclude the lagged patent output from sector i itself; doing so ensures the coefficient β_1 in regression (25) is not driven by serially correlated shocks to sectoral patent output. Second, theoretically, the knowledge aggregator in (9) features exponential decay of past patents' effects, yet our empirical construction (24) features a discrete cutoff window for $\tau \leq 10$ years. We choose this to be agnostic about the parameter λ; later we also nonparametrically estimate the effect at different time lags.

Table 2, column (1) presents the results of regression (25). 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—$\text{Knowledge}_{it}^{Up}$, or $\ln \chi_{it}$—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.7, 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 for our proposed mechanism, 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 (25) even without cross-sector knowledge spillovers. This is

[4] Our empirical results are robust to alternative values of τ (see Online Appendix Table A.9 for $\tau = 5$ and 20).

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)
$Knowledge_{it}^{Up}$	0.586***	0.600***	0.508***	0.679**	0.802***	0.830***	0.743***	0.974***
	(0.180)	(0.205)	(0.174)	(0.266)	(0.202)	(0.218)	(0.196)	(0.279)
$ln(R\&D)_{i,t-1}$	0.275***	0.274***	0.279***	0.269***	0.258***	0.256***	0.261***	0.174**
	(0.063)	(0.062)	(0.060)	(0.070)	(0.086)	(0.086)	(0.086)	(0.082)
$Knowledge_{it}^{Down}$		-0.029				-0.058		
		(0.157)				(0.098)		
$Knowledge_{it}^{Up,IO}$			0.363**				0.268	
			(0.173)				(0.205)	
Specification	OLS	OLS	OLS	IV 2nd Stage	OLS	OLS	OLS	IV 2nd Stage
IV 1st Stage F-statistics				465.9				465.9
R^2	0.915	0.915	0.917	0.152	0.901	0.901	0.902	0.099
No. of Sectors	94	94	94	94	94	94	94	94
No. of Obs	1847	1847	1847	1113	1847	1847	1847	1113
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. To measure innovation production (Y), $\ln$ (Patents) is the number of new patents in each sector-year, and $\ln$ (cites) is the (log-)number of total future citations (up to the end of our sample) received by new patents in each sector-year, which is equivalent to citation-weighted new patent counts in that sector-year. The key variable of interest, $\text{Knowledge}_{it}^{Up}$, is the knowledge from upstream, defined in (24). Lagged sectoral R&D expenses and sector and year fixed effects are included as controls. Columns (2) and (6) include downstream knowledge as a control. Columns (3) and (7) include knowledge accumulated from upstream sectors in the production network as a control. Columns (4) and (8) present 2SLS estimation where $\text{Knowledge}_{it}^{Up}$ is instrumented using tax-induced variations in the user cost of R&D capital. Standard errors in parentheses are clustered at the sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

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:

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

That is, $\text{Knowledge}_{it}^{Down}$ aggregates the (log-)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, our theory implies the following asymmetry: while the upstream aggregator $\text{Knowledge}_{it}^{Up}$ should positively predict subsequent patent output in sector i, the downstream aggregator $\text{Knowledge}_{it}^{Down}$ 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.

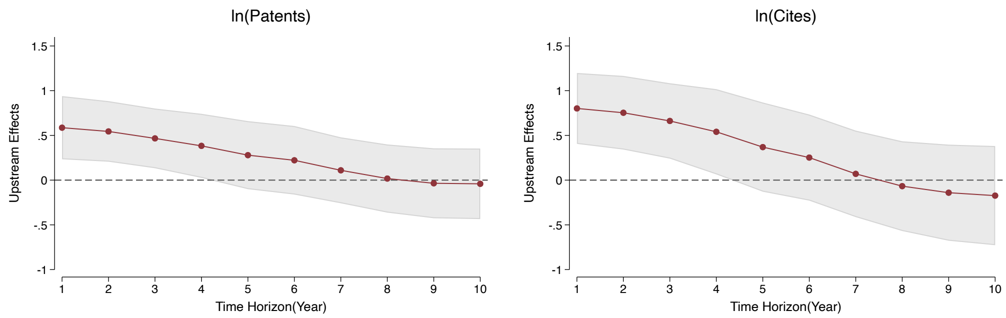
Another related concern is that common shocks operate not through technological linkages but through input-output (IO) linkages. To address this, we construct the aggregator $\text{Knowledge}_{it}^{Up,IO}$ similarly to $\text{Knowledge}_{it}^{Up}$, but using patents from other sectors weighted not by the innovation network, as in (24), 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, $\text{Knowledge}_{it}^{Up,IO}$ does not significantly predict sector i's innovation in these specifications. We find the coefficient on $\text{Knowledge}_{it}^{Up,IO}$ is sometimes significant when we omit the main variable $\text{Knowledge}_{it}^{Up}$ from the regressions, but the effects are dwarfed by the spillover effects through the innovation network (see Table A.8 in the Online Appendix). 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, 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 (24) 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.

Finally, we revisit the dynamic prediction of our key law of motion (9), that upstream knowledge from the more distant past has less effect on patent output. To explore this, we perform our baseline regression (25) using lagged versions of the upstream knowledge aggregator on the right-hand side. We plot the coefficients and confidence intervals in Figure (4), where the one-year lag corresponds to the baseline estimates in columns (1) and (5) in Table 2. 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.

Figure 4. 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.

5.2.2 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:

$$\text{Knowledge}_{mit}^{Up} \equiv \sum_{cj \neq mi} \sum_{\tau=1}^{10} \frac{Cites_{mi \to cj, t-\tau}}{\sum_{c'=1}^{N} \sum_{k=1}^{K} Cites_{mi \to c'k,, t-\tau}} \ln n_{cj,t-\tau}. \tag{26}$$

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:

$$\ln n_{mit} = \beta_1 \times \text{Knowledge}_{mit}^{Up} + \beta_2 \times \ln R\&D_{mi,t-1} + \xi_{mi} + \xi_{mt} + \xi_{it} + \text{controls}_{mit} + \varepsilon_{ict}. \tag{27}$$

We include a saturated set of fixed effects. The country-industry fixed effect ξ_{mi} accounts for any time-invariant unobserved heterogeneity in patent output (e.g., IT industries in the U.S. and France have different patent productivity); the country-year fixed effects ξ_{mt} control for time-varying country-level shocks (e.g., patent productivity, business cycles) that are common across industries; and the industry-year fixed effects ξ_{it} account for time-varying global shocks to patent output that are common within industries and across countries.

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 knowledge spillovers are stronger across sectors within the U.S. than they are across countries.

To rule out common shocks to technological and input-output clusters, we again—similar to our closed economy tests—construct and control for aggregators $\text{Knowledge}_{mit}^{Down}$ to capture knowledge from downstream and $\text{Knowledge}_{mit}^{Up,IO}$ to capture potential knowledge spillovers through the input-output network. Columns (2), (3), (6), and (7) show that coefficients on these controls are insignificant, and our coefficients on $\text{Knowledge}_{mit}^{Up}$ do not materially change when these controls are added. We also produce the global versions of the IV-results in columns (4) and (8). Overall, these results validate our mechanism of knowledge spillovers through the international innovation network.

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

$Y=$	ln(Patents)				ln(Cites)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Knowledge_{mit}^{Up}$	0.157***	0.199***	0.154***	0.202*	0.348***	0.424***	0.345***	0.405***
	(0.057)	(0.061)	(0.057)	(0.113)	(0.083)	(0.089)	(0.084)	(0.147)
$ln(R\&D)_{mi,t-1}$	0.033***	0.034***	0.032***	0.066***	0.046***	0.046***	0.046***	0.072***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)	(0.014)	(0.022)
$Knowledge_{mit}^{Down}$		-0.091**				-0.167***		
		(0.043)				(0.059)		
$Knowledge_{mit}^{Up,IO}$			0.079				-0.038	
			(0.064)				(0.070)	
Specification	OLS	OLS	OLS	IV 2nd Stage	OLS	OLS	OLS	IV 2nd Stage
IV 1st Stage F-statistics				146.2				146.2
R^2	0.969	0.969	0.969	0.040	0.944	0.944	0.945	0.031
No. of Country x Sectors	564	564	550	280	564	564	550	280
No. of Obs	10,552	10,552	10,318	4,467	10,552	10,552	10,318	44,67
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. To measure innovation production (Y), we use the number of patents and total number of citations. The key variable of interest, $\text{Knowledge}_{it}^{Up}$, is the knowledge from upstream, defined in (26). Fixed effects at the country-sector, country-year, and sector-year levels are included as controls. Columns (2) and (6) include downstream knowledge as a control. Columns (3) and (7) include knowledge accumulated from upstream sectors in the production network as a control. Columns (4) and (8) present 2SLS estimation where $\text{Knowledge}_{it}^{Up}$ is instrumented using tax-induced variations in the user cost of R&D capital. Standard errors in parentheses are clustered at the country-sector level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

6 Application: R&D Resource Allocations in the Data

This section uses our model to evaluate the allocation of R&D resources in the data. We compute the unilaterally optimal allocation of R&D resources across sectors for each country and year in our sample. We show that on average, sectors that 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 cost of R&D misallocation for each country, and we show R&D misallocation is less severe in economies where innovation activities are concentrated in a small number of innovative firms.

6.1 Optimal R&D Allocations

For each country and year, we calculate the unilaterally optimal cross-sector allocation of R&D resources $\boldsymbol{\gamma}$, using Proposition 8:

$$\boldsymbol{\gamma}' = \xi^{-1}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}\circ\boldsymbol{X}}{1+\rho/\lambda}\right)^{-1}, \tag{28}$$

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. 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.

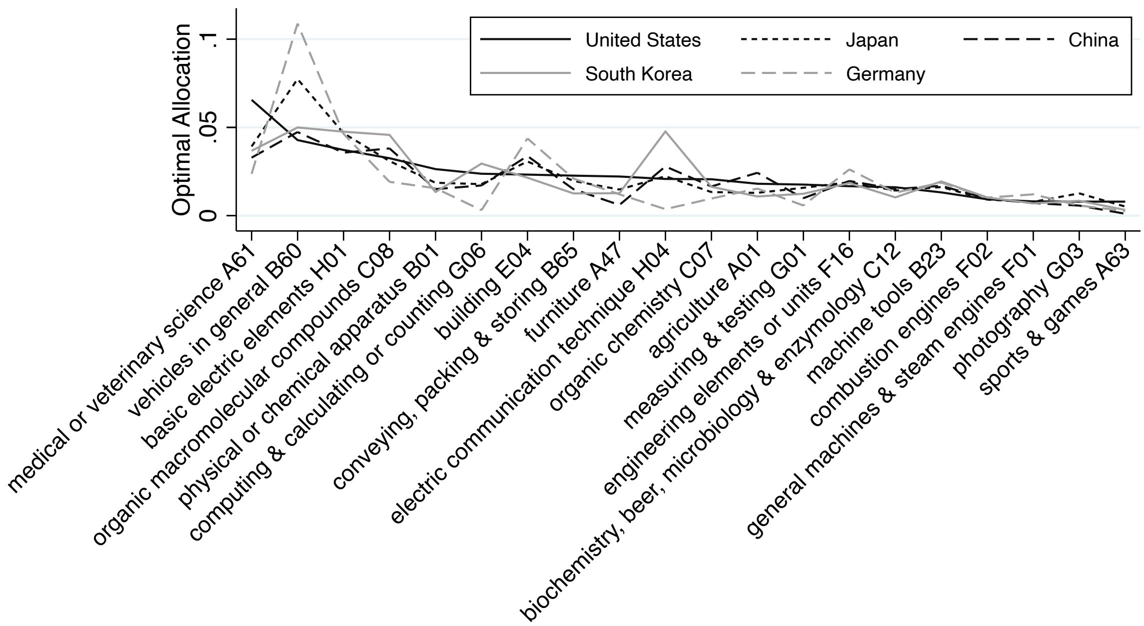
To implement the formula in (28), we need to specify the discount rate relative to the innovation step size, ρ/λ. We set discount rate $\rho = 5\%$ and step size $\lambda = 0.05$, in line with recent literature on quality-ladder models (e.g., Akcigit and Ates, 2022). Unless otherwise noted, we use data averaged across 2010–2014. Qualitatively, our cross-country analysis is not sensitive to the value of ρ/λ;[5] for our quantitative exercises, we report alternative values of ρ/λ in the Online Appendix as sensitivity checks. The Online Appendix also reports additional results using data from other sample periods.

Figure 5 plots the optimal R&D allocation $\boldsymbol{\gamma}$ for the five economies that produced the most 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.,

[5] As discussed previously, an increase in ρ/λ has the same implication for 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.

Figure 5. 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 $(1+\rho/\lambda)^{-1} = 0.5$. Sectors are sorted by the optimal allocation for the U.S.,γ_{US}.

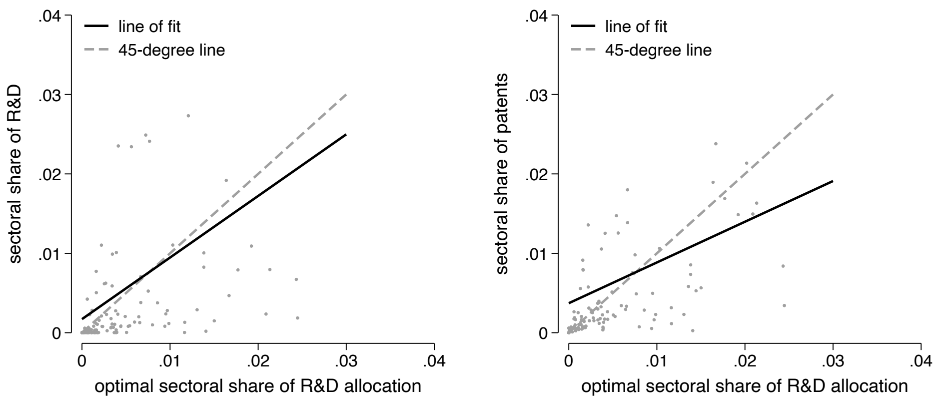
semiconductors), and computing devices (G06). The top 10 IPCs (out of 131) should receive about a third of total U.S. R&D resources. The correlation between optimal U.S. R&D allocation and the innovation centrality $\boldsymbol{a}$ is 0.77. The correlation is high because the U.S. has a self-contained 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 5 also reveals cross-country variations in optimal R&D allocations. The variation originates from heterogeneity in the production structure ($\boldsymbol{\beta}$) and R&D's degree of foreign reliance ($\boldsymbol{X}$). 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).

6.2 Innovation Allocations in the Data

We first present our model's ability to fit R&D resource allocations in the U.S. The left panel of Figure 6 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 0.78 (t-statistic 5.09), 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 6, 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 0.51 (t-statistic 3.85). 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 cost of this misallocation in Section 6.3 below.[6]

Figure 6. U.S. Actual R&D Allocation vs. Optimal Allocation$\boldsymbol{\gamma}_{US}$



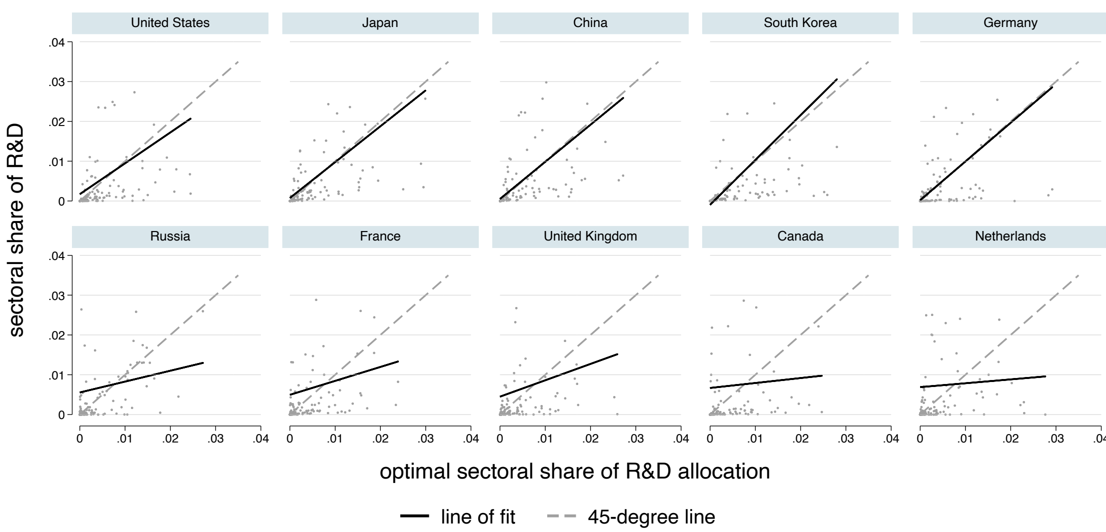
Notes. This figure shows scatter plots of real-world sectoral R&D expenditure shares (left panel) and patent output shares (right panel) against optimal R&D allocation shares, γ_{US}, for the U.S. in 2010-2014. The solid line is the linear fit; the dashed line is the 45-degree line.

There is substantial cross-country heterogeneity in R&D resource allocations. Figure 7 shows scatter plots of sectoral R&D expenditure shares against unilaterally optimal R&D allocations for ten countries that filed the most patents in 2010–2014. Sectoral R&D expenditure correlates strongly with 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 6 suggests that on average, more resources need to be reallocated to achieve optimality in the five economies

[6]A potential concern is that Figure 6 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 centrality of the innovation network. 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 innovation 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.

in the bottom row. These results are robust to using patent output as a proxy of cross-sector R&D resource allocation (Figure A.8 in the Online Appendix), suggesting that our results are not driven by the coverage and quality of our R&D expenditure variable. In Figure A.9 in the Online Appendix, we plot the alignment between real and optimal allocation in the years 2000 and 2005. We find very similar patterns, except for China, for which the fit improved significantly from 2000 to 2010.

Figure 7. Actual R&D Allocation vs. Optimal Allocation Across Countries



Notes. This figure shows scatter plots of sectoral patent output (as a fraction of total patent output in each country) 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, outliers—sectors that account for >5% of national patent output—are not shown in the scatter plots, but all sectors are used when constructing the linear fit.

6.3 Welfare Cost of R&D Misallocation

We now quantify the welfare cost of R&D misallocation. Under our baseline parameters $\rho = \lambda = 0.05$, the welfare cost formula in Proposition 9 is:

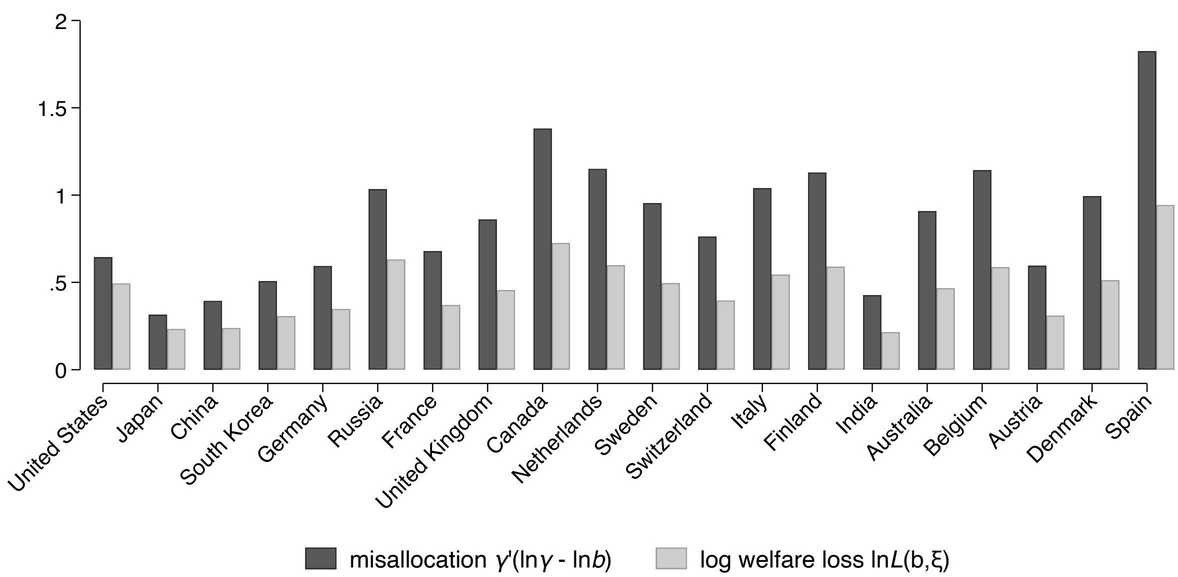
$$\ln \mathcal{L}(\boldsymbol{b}, \xi) = \underbrace{\xi}_{\text{self sufficiency}} \times \underbrace{\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})}_{\text{R\&D misallocation}}, \tag{29}$$

where $\ln \mathcal{L}(\boldsymbol{b}, \xi)$ is the consumption-equivalent welfare cost in logs, $\boldsymbol{b}$ is the empirical R&D allocation vector, $\boldsymbol{\gamma}$ is the optimal R&D allocation vector, and $\xi \equiv \frac{\rho}{\rho+\lambda}\left(\boldsymbol{I} - \frac{\boldsymbol{\Omega} \circ \boldsymbol{X}}{1+\rho/\lambda}\right)^{-1} \mathbf{1}$ is the scalar measure of R&D self-sufficiency, which increases in domestic citation shares x_{ij}.

Equation (29) implies a natural decomposition when comparing the welfare cost of R&D misallocation across countries: misallocation ($\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})$) and self-sufficiency ($\xi$). Holding an economy's R&D self-sufficiency constant, worse misallocation leads to larger welfare costs, and as an economy relies more on foreign knowledge spillovers (lower ξ), domestic R&D misallocation becomes less consequential for consumer welfare.

We focus on R&D misallocation across top 50 IPC classes in terms of global patenting activities. These IPCs jointly account for 92% of all patents in our data and 90% of all R&D expenses. Focusing on the top innovative sectors ensures that our results are not driven by misallocation in technologies with few patents, as our theory may be less applicable to these sectors.[7] Online Appendix E.4 provides additional results using all IPCs in the economy.

Figure 8. R&D Misallocation and Welfare Cost Across Countries



Notes. This table shows the level of R&D misallocation (dark bars) and associated welfare cost (light grey bars) across 20 innovative countries with the highest patent outputs in our sample, using 2010-2014 data. The calculation focuses on misallocation in top 50 IPC classes by total patents.

Figure 8 shows the welfare cost of R&D misallocation for the 20 economies that filed the most patents during 2010–2014. The dark bars represent the misallocation term. Among high-income countries, Japan has the lowest R&D misallocation, followed by South Korea and the U.S. China and India, both fast-growing developing economies, have less R&D misallocation than most high-income economies in Europe. The light grey bars represent the welfare cost of R&D

[7]Our results therefore use the welfare formula for when the planner can reallocate resources across a subset of sectors; see Section 2.5 and Online Appendix Section B.5 for the closed and open economy versions, respectively.

misallocation, taking into account each country's R&D self-sufficiency. For the two economies with self-contained innovation networks, namely the U.S. and Japan, ξ is closer to one, so the overall welfare costs (dark bars) are closer in magnitude to the misallocation terms (grey bars). By contrast, welfare costs are significantly 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.

In terms of magnitudes, adopting the optimal allocations in Japan could lead to welfare improvements equivalent to raising consumption along the entire path by $(\exp(0.234) - 1 =)$ 26%, the lowest among countries in our sample. The potential welfare gains for the U.S. are $(\exp(0.495) - 1 =)$ 64% in consumption-equivalent terms,[8] which is above the average in our sample. Spain has the highest welfare cost of R&D misallocation, equivalent to leaving $(\exp(1.051) - 1 =)$ 186% consumption gains on the table.

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 cost 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 Japan's Level of R&D Allocative Efficiency

	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
2000	19.09	-	43.79	36.11	33.93	39.32	27.09	30.44	146.40	47.41
2005	24.25	-	33.31	14.24	25.64	64.07	19.77	40.35	71.89	67.50
2010	28.76	-	4.94	12.36	17.78	55.20	22.01	33.57	75.21	54.43

	Sweden	Switzerland	Italy	Finland	India	Australia	Belgium	Austria	Denmark	Spain
2000	55.36	17.79	28.64	32.16	4.20	23.10	94.90	8.16	56.32	25.95
2005	33.30	19.12	72.25	38.00	5.45	32.31	71.00	11.48	47.73	28.05
2010	39.48	26.26	46.27	53.08	5.81	35.69	53.04	15.72	41.95	118.34

Notes. This table shows the consumption-equivalent welfare gains when each economy moves to Japan's level of R&D allocative efficiency. Specifically, this table shows the welfare gain of moving from country-specific levels of misallocation, captured by $\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b})$, to Japan's level $\boldsymbol{\gamma}'_{JP}(\ln\boldsymbol{\gamma}_{JP} - \ln\boldsymbol{b}_{JP})$ in the corresponding year. The calculation focuses on improving allocation in top 50 IPC classes by total patents.

While efficient R&D allocation is country-specific, the distance from each country's R&D to efficiency can be compared across countries. For each country in our sample, we calculate the

[8] 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 the actual R&D expenditures in the U.S. to be marginally closer to optimal under that specification, and the potential welfare gains are correspondingly smaller (126% increase in consumption).

consumption-equivalent welfare gains for adopting Japan's level of R&D allocative efficiency by subtracting Japan's misallocation term from each country's own term before applying formula (29). Table 4 shows that moving to "Japan's efficiency" in 2010 would generate consumption-equivalent welfare gains of 28.8% in the U.S., 4.9% in China, 12.4% in South Korea, and 17.8% in Germany.

In Online Appendix E.4, we present additional results on the welfare cost of R&D misallocations: we reproduce Figure 8 using sectoral shares of patent output (instead of R&D expenditure shares) as the real allocation variables; we reproduce Figure 8 and Table 4 using all IPCs (instead of top 50); we also present over time misallocation patterns for the 10 countries with the most patents and reproduce some of this section's results using alternative values of ρ and λ.

6.4 How Are R&D Resources Misallocated in the U.S.?

We here provide some descriptive evidence for how R&D resources are misallocated in the U.S. Figure 9 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 81% of R&D expenditures in the U.S. Though providing a full set of policy recommendations on R&D allocations is beyond the scope of this paper, this figure reveals several noteworthy facts. Electric communication technique (H04; e.g., telephonic communication, wireless communication), which ranks 4th in centrality (Figure 2) and 10th in the optimal allocation γ (Figure 5), is over-invested. Meanwhile, within the same broad IPC class H, the more central and fundamental class Basic Electric Elements (H01; e.g., semiconductor devices), is underinvested. This supports recent U.S. initiatives to invest in the semiconductor industry.[9] Another observation is that the underinvested group (right end of the graph) over-represents IPC classes related to technologies often termed as "green innovation" that can help reduce pollution and the negative consequences of resource exploitation. For instance, the most underinvested IPC class, B01, covers subclasses on waste management, alternative energy production, and environmental management.[10]

We can leverages Proposition 5 to provide guidance on allocation within a specific technology or sector.[11] Figure 10 demonstrates R&D misallocations *within* broad technology classes (1-digit IPC) for the U.S. economy. Each of the eight 1-digit IPC categories ("A" through "H") 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

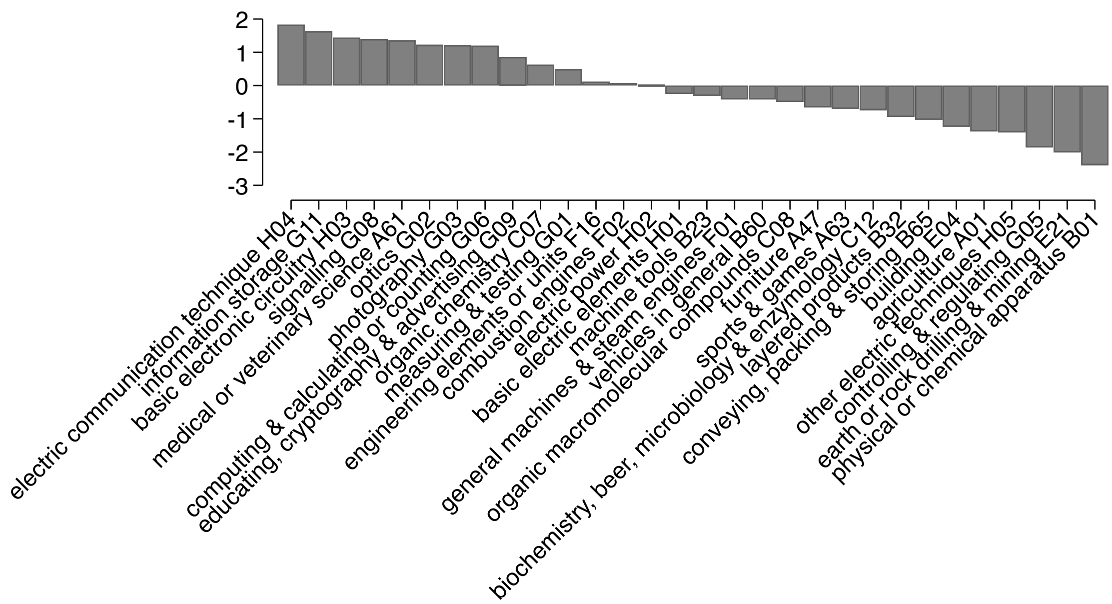
[9]For example, the CHIPS For America Act, which was passed in January 2021, will unlock more than 50 billion USD to develop the domestic semiconductor industry.

[10]For more references on identifying green innovation-related IPC classes, see Cohen, Gurun, and Nguyen (2020).

[11]The open economy version of the proposition is in Section B.5 of the Online Appendix.

the 1-digit IPC category. For the ease of visualization, we show the five 3-digit IPC classes with the highest R&D expenditure levels within each panel; the 40 IPC classes represented across all eight panels in Figure 10 account for 88% of U.S. R&D. We find that in IPC category A (i.e., human necessities), medical or veterinary science (A61) is over-allocated with R&D relative to agriculture (A01) and foods (A23); in IPC category C (chemistry; metallurgy), organic chemistry (C07) is over-allocated relative to petroleum, gas, or coke industries (C10); in category E (fixed construction), locks, keys, and safes (E05) is over-allocated relative to mining (E21).

Figure 9. U.S. R&D Misallocation in the Top 30 Innovative IPC Classes

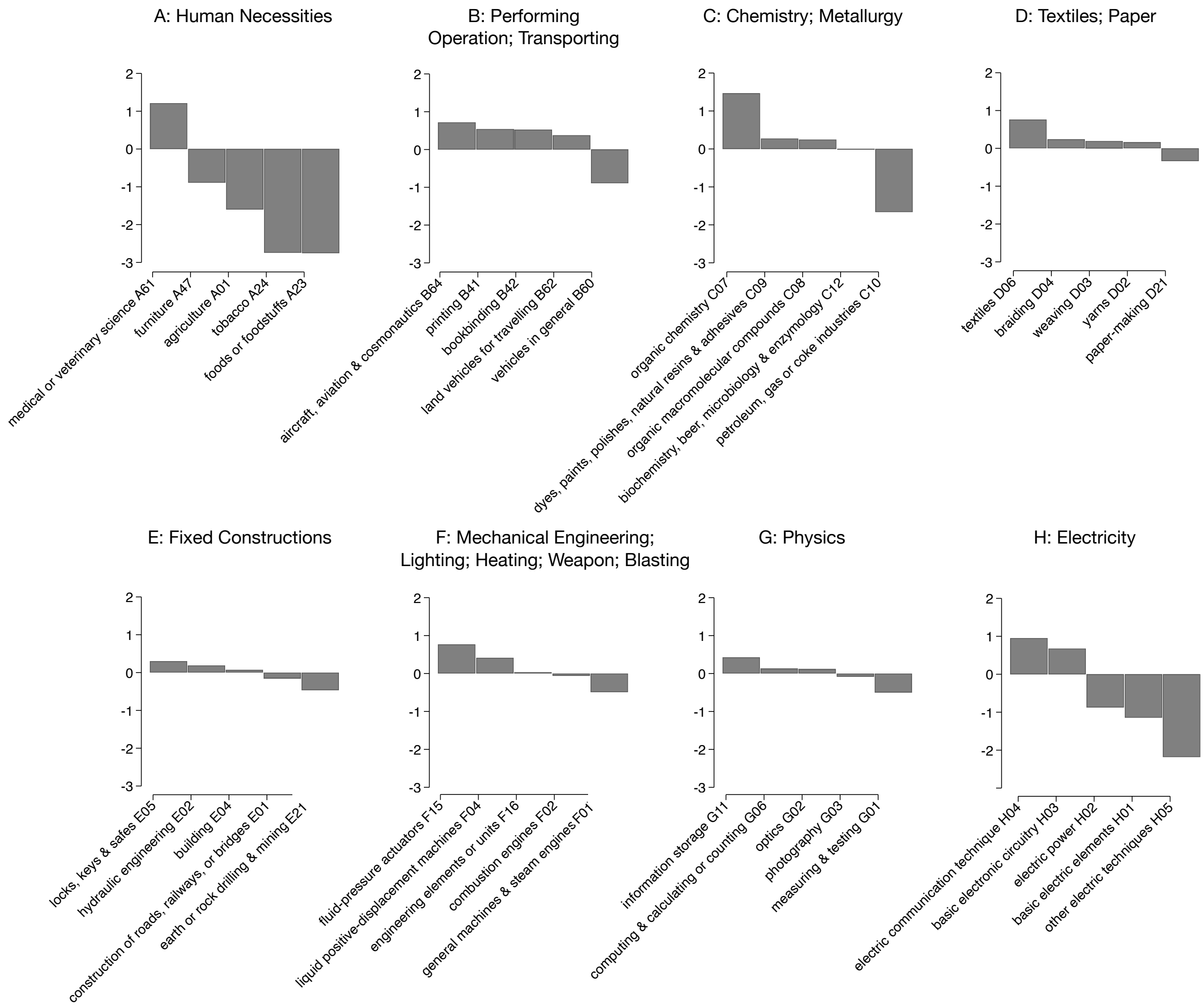


Notes. This figure plots the level of misallocation of the top 30 innovative IPC classes, ranked using total patent output. The level of misallocation is calculated as $\ln \boldsymbol{b} - \ln \boldsymbol{\gamma}$. Positive bars (left end) mean over-investment, and negative bars mean underinvestment.

6.5 Innovation Hubs

What explains cross-country differences in R&D misallocation? 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 ex-

Figure 10. U.S. R&D Misallocation within 1-digit IPC Classes



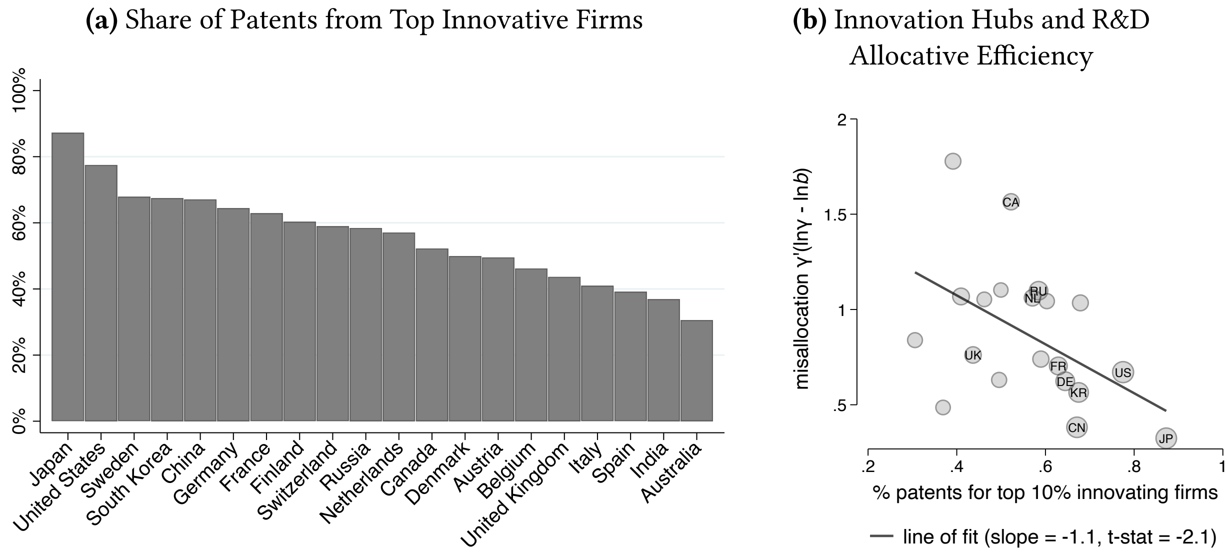
Notes. This figure shows U.S. R&D misallocation across 3-digit IPC classes within each 1-digit IPC class. For each 1-digit IPC class, $\ln \boldsymbol{b} - \ln \boldsymbol{\gamma}$ is shown for the top five 3-digit IPC classes ranked by R&D expenditures.

amples 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 11, Panel (a) shows the share of patents in 2010–2014 that are filed by the top 10% of innovating firms in each country. 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 11 plots the misallocation measure (i.e., the distance to efficient R&D allo-

Figure 11. 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).

cation, $\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b})$) against the share of patents accounted for by the top 10% of innovating firms. We find a strongly negative relationship (slope -1.1, 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.

7 Conclusion

A large literature has highlighted sources of inefficiency in innovation, specifically regarding under- or over-investment in R&D. 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 upstream in the innovation network, but the incentive is muted in open economies that rely on foreign knowledge spillovers. We show the relative entropy of actual R&D allocation from the optimal allocation is a sufficient statistic for the welfare cost of R&D misallocation.

To 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 consumption, production, 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, significant misallocation remains, leading to sizable welfare costs. Adopting optimal 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 55% 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. and S. Ates (2022): "What Happened to U.S. Business Dynamism?" *Journal of Political Economy.*

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 (forthcoming): "Knowledge Diffusion, Trade and Innovation across Countries and Sectors," *AEJ: Macroeconomics.*

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.*

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.*

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.*

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. I. (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 (2021a): "Dynamic Spatial General Equilibrium," *Working Paper.*

——— (2021b): "International Friends and Enemies," *Working Paper.*

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 (2021): "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.*
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.*
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.
Yang, Y. and W. Zhu (2020): "Innovation Network, Production Network, and Business Cycles," *Working paper.*

Online Appendix (For Online Publication Only)

Table of Contents

A Proofs

A.1 Proof of Lemma 1: Optimal Labor Allocation in a Closed Economy

The planner's problem is

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

subject to constraints (3), (5), (6), and (7). Substituting using (3) and (5), the objective can be re-written as

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

The FOC with regard to $\ell_{it}(\nu)$ gives:

$$\frac{\beta_i}{\ell_{it}(\nu')} = \frac{\beta_i}{\ell_{it}(\nu)} = \frac{\beta_j}{\ell_{jt}(\nu)} \quad \Rightarrow \quad \ell_{it}(\nu) = \ell_{it}, \quad \frac{\ell_{jt}(\nu)}{\ell_{jt}(\nu)} = \frac{\beta_j}{\beta_i}.$$

Therefore, for all t, $\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell}$ for each sector i and variety ν.

A.2 Proof of Proposition 1: Optimal R&D Allocation in a Closed Economy

The social planner's problem is

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

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

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

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

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

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):

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

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

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

as desired.

A.3 Proof of Lemma 2: Growth Rate Along a Closed Economy BGP

The law of motion for stock vector is

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

Consider a BGP in which R&D allocation shares follow the vector $\boldsymbol{b}$. Suppose the knowledge stock in every sector grows at the rate $g(\boldsymbol{b})$, then

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

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

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

The third equation is based on the properties of the innovation centrality vector: $\boldsymbol{a}'=\boldsymbol{a}'\boldsymbol{\Omega}$ and $\sum_{i=1}^{K}a_i=1$.

A.4 Proof of Proposition 2

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

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

Taking the limit as $\rho/\lambda\to 0$, $\boldsymbol{\gamma}'\left(\boldsymbol{I}-\boldsymbol{\Omega}\right)\to\mathbf{0}$ 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 Allocations in a Closed Economy

Given the law of motion for sectoral knowledge stock

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

we can solve for the evolution of knowledge stock in closed form as a function of R&D allocation $\boldsymbol{b}$:

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

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

$$\begin{aligned} V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \widetilde{\boldsymbol{b}}\right) - V\left(\boldsymbol{q}_0; \{\boldsymbol{\ell}_t\}, \boldsymbol{b}\right) &= \int_0^\infty e^{-\rho t} \left[\ln \boldsymbol{q}_t\left(\widetilde{\boldsymbol{b}}\right) - \ln \boldsymbol{q}_t\left(\boldsymbol{b}\right)\right] \,\mathrm{d}t \\ &= \lambda \int_0^\infty e^{-\rho t} \left[\int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega})(t-s)} \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b}\right) \,\mathrm{d}s\right] \,\mathrm{d}t. \end{aligned}$$

Let $\boldsymbol{\Omega} \equiv \boldsymbol{U\Lambda V}$ denote the eigendecomposition of $\boldsymbol{\Omega}$, where the k-th column of $\boldsymbol{U}$ ($\boldsymbol{u}_k$) and k-th row of $\boldsymbol{V} \equiv \boldsymbol{U}^{-1}$ ($\boldsymbol{v}_k'$) are respectively the k-th right- and left-eigenvectors of $\boldsymbol{\Omega}$, with the associated eigenvalue ψ_k being the k-th entry on the diagonal of $\boldsymbol{\Lambda}$, which is a diagonal matrix:

$$\boldsymbol{\Omega u}_k = \psi_k \boldsymbol{u}_k, \quad \boldsymbol{v}_k' \boldsymbol{\Omega} = \psi_k \boldsymbol{v}_k'.$$

We can re-write the difference in consumer welfare as

$$
\begin{aligned}
& \lambda \boldsymbol{\beta}' \int_0^\infty e^{-\rho t} \left[\int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Omega})(t-s)} \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \, \mathrm{d}s \right] \, \mathrm{d}t \\
= \; & \lambda \boldsymbol{\beta}' \sum_{k=1}^K \int_0^\infty e^{-\rho t} \int_0^t e^{-\lambda(1-\psi_k)(t-s)} \, \mathrm{d}s \, \mathrm{d}t \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \boldsymbol{\beta}' \sum_{k=1}^K \int_0^\infty e^{-\rho t} \left[e^{-\lambda(1-\psi_k)t} - 1 \right] \, \mathrm{d}t \, (1-\psi_k)^{-1} \, \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \boldsymbol{\beta}' \sum_{k=1}^K \int_0^\infty \left[e^{-(\rho+\lambda-\lambda\psi_k)t} - e^{-\rho t} \right] \, \mathrm{d}t \, (1-\psi_k)^{-1} \, \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \boldsymbol{\beta}' \sum_{k=1}^K \left[\frac{1}{\rho+\lambda-\lambda\psi_k} - \frac{1}{\rho} \right] \frac{1}{1-\psi_k} \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \boldsymbol{\beta}' \sum_{k=1}^K \left[\frac{\lambda}{\rho\,(\rho+\lambda-\lambda\psi_k)} \right] \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \frac{\lambda}{\rho^2} \frac{\rho}{\rho+\lambda} \boldsymbol{\beta}' \sum_{k=1}^K \left[\frac{1}{1-\frac{1}{1+\rho/\lambda}\psi_k} \right] \boldsymbol{u}_k \boldsymbol{v}_k' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right) \\
= \; & \frac{\lambda}{\rho^2} \boldsymbol{\gamma}' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right),
\end{aligned}
$$

as desired.

A.6 Proof of Proposition 4: Consumption-Equivalent Welfare Cost of R&D Misallocation in a Closed Economy

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: Constrained Optimal R&D Allocations & Welfare

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. From Proposition 3, $\boldsymbol{\gamma}^{\mathcal{K}}$ is 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 $\gamma_i^{\mathcal{K}} = \frac{\gamma_i}{\sum_{j \in \mathcal{K}} \gamma_j}$. Again by Proposition 3, the welfare cost of misallocation is

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

which simplifies to the formula in Proposition 5.

A.8 Proof of Proposition 6: Allocations in the Decentralized Equilibrium

We normalize the consumer price index to one for all times t. The consumer at each time t spends a constant fraction β_i of their income on sectoral composite good i, with

$$p_{it}y_{it} = \beta_i y_t \quad \text{for all } i, t. \tag{A5}$$

The sectoral composite aggregator (3) 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 $(1+\lambda)$, we derive the profits in each sector i as

$$\pi_{it}(\nu) = \frac{\lambda}{1+\lambda}\beta_i y_t \quad \text{for all } i, t, \nu. \tag{A6}$$

Because all varieties have identical markups, the worker allocation is identical across varieties within each sector. The total labor cost in each sector i is $\frac{1}{1+\lambda}\beta_i y_t$ and is thus also proportional to the consumption shares β_i:

$$\ell_{it}(\nu) = \ell_{it} = \beta_i \bar{\ell} \quad \text{for all } i, t, \nu. \tag{A7}$$

Along the BGP, a monopolist in each sector has the same Poisson rate to be replaced by an innovating entrant. Let δ denote that replacement rate; the value of a monopolistic firm is thus

$$v_{it} \equiv \int_t^\infty e^{-(r_s+\delta)(s-t)}\pi_{is}\,\mathrm{d}s, \tag{A8}$$

where r_s is the interest rate at time s and g_s is the growth rate of aggregate consumption y_s. 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

$$v_{it}/v_{jt} = \beta_i/\beta_j \quad \text{for all } i, j, t. \tag{A9}$$

Entrants hire scientists to conduct research in order to become future monopolists. The marginal value from an additional scientist ($v_{it} \times \partial \ln(n_{it}/q_{it})/\partial s_{it}$) must be equalized across sectors. Substituting n_{it} using the innovation production function (5) and v_{it}/v_{jt} using equation

(A9), we obtain that scientist allocation must also follow the consumption share, that is, $s_{it}/\bar{s} = \beta_i$ for all t, as desired.

A.9 Proof of Proposition 7: Innovation Policy that Decentralizes the Optimal Allocations

Suppose sector i pays proportional R&D tax t_i. The marginal value of R&D spending in sector i is $\frac{v_{it}}{1+t_i}\frac{\partial \ln(n_{it}/q_{it})}{\partial s_i}/w_t^s = \frac{v_{it}}{s_{it}(1+t_i)}/w_t^s$, where v_{it} is the value of a firm in sector i as in equation (A8), and w_t^s is the wage rate of scientists. Because $v_{it}/v_{jt} = \beta_i/\beta_j$ for all i, j, we know

$$\frac{v_{it}}{s_{it}\left(1+t_i\right)}\Bigg/\frac{v_{jt}}{s_{jt}\left(1+t_j\right)} = \frac{\beta_i s_{jt}\left(1+t_j\right)}{\beta_j s_{it}\left(1+t_i\right)}.$$

When $1+t_i \propto \frac{\beta_i}{\gamma_i}$, it must be the case that $\frac{\gamma_i}{\gamma_j} = \frac{s_i}{s_j}$. Thus, the planner can decentralize the optimal allocation by setting R&D taxes to be $1+t_i \propto \frac{\beta_i}{\gamma_i}$, with the appropriate lump-sum tax levied on the consumer to balance the budget.

A.10 Proof of Proposition 8: Optimal Allocations in an Open Economy

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

$$\bar{\mathcal{C}}^*\left(y_t, p_t^f\right) \equiv \max_{c_t^d, c_t^f} \mathcal{C}\left(c_t^d, c_t^f\right) \quad \text{s.t. } y_t - c_t^d = p_t^f c_t^f. \tag{A10}$$

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 $\boldsymbol{b}_t$:

$$\ln \boldsymbol{q}_t = e^{\lambda(\boldsymbol{\Theta}-\boldsymbol{I})t}\left[\ln \boldsymbol{q}_0 + \lambda \int_0^t e^{-\lambda(\boldsymbol{\Theta}-\boldsymbol{I})s}\left(\left(\boldsymbol{\Omega}-\boldsymbol{\Theta}\right)\ln \boldsymbol{q}_s^f + \ln \boldsymbol{\eta} + \ln \bar{s} + \ln \boldsymbol{b}_s\right) \mathrm{d}s\right]. \tag{A11}$$

The optimal R&D allocation is

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

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

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

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

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

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

as desired.

A.11 Proof of Proposition 9: Welfare Cost of R&D Misallocation in an Open Economy

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 (A10). Following the proof to Proposition 8, $\bar{\mathcal{C}}^*\left(y_t, p_t^f\right) = y_t \mathcal{C}^*\left(p_t^f\right)$; 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 $\boldsymbol{\beta}' \ln \boldsymbol{q}_t$, 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 (A11), we can re-write the welfare differences as

$$V(\boldsymbol{\gamma}) - V(\boldsymbol{b}) = \int_0^\infty e^{-\rho t}\left[\lambda \int_0^t e^{-\lambda(\boldsymbol{I}-\boldsymbol{\Theta})(t-s)}\,\mathrm{d}s\right]\mathrm{d}t\,(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}),$$

where $\boldsymbol{\Theta} \equiv \boldsymbol{\Omega} \circ \boldsymbol{X}$. To simplify the integral we follow the proof to Proposition 3 and undertake an eigendecomposition to $\boldsymbol{\Theta}$. Let $\boldsymbol{u}_k$, $\boldsymbol{v}_k'$, and ψ_k denote the k-th right-eigenvector, left-eigenvector, and eigenvalue, respectively; then

$$\begin{aligned}
V(\boldsymbol{\gamma}) - V(\boldsymbol{b}) &= \lambda\boldsymbol{\beta}' \sum_{k=1}^{K} \boldsymbol{u}_k \boldsymbol{v}_k' \int_0^\infty e^{-\rho t}\left[\int_0^t e^{-\lambda(1-\psi_k)(t-s)}\,\mathrm{d}s\right]\mathrm{d}t\,(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \boldsymbol{\beta}' \sum_{k=1}^{K} \frac{1}{1-\psi_k}\boldsymbol{u}_k \boldsymbol{v}_k' \int_0^\infty e^{-\rho t}\left[\left(e^{-\lambda(1-\psi_k)t} - 1\right)\right]\mathrm{d}t\,(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \boldsymbol{\beta}' \sum_{k=1}^{K} \frac{\lambda}{\rho(\rho+\lambda-\lambda\psi_k)}\boldsymbol{u}_k \boldsymbol{v}_k' (\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \frac{\lambda}{\rho^2}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \sum_{k=1}^{K} \left[\frac{1}{1-\frac{1}{1+\rho/\lambda}\psi_k}\right]\boldsymbol{u}_k \boldsymbol{v}_k' (\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \frac{\lambda}{\rho^2}\frac{\rho}{\rho+\lambda}\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1+\rho/\lambda}\right)^{-1} (\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \frac{\lambda}{\rho^2}\frac{\rho}{\rho+\lambda}\left(\boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1+\rho/\lambda}\right)^{-1}\boldsymbol{1}\right) \frac{\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1+\rho/\lambda}\right)^{-1}}{\boldsymbol{\beta}'\left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1+\rho/\lambda}\right)^{-1}\boldsymbol{1}} (\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}) \\
&= \frac{\lambda}{\rho^2}\xi\boldsymbol{\gamma}' (\ln\boldsymbol{\gamma} - \ln\boldsymbol{b}).
\end{aligned} \tag{A12}$$

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 cost of R&D misallocation is thus

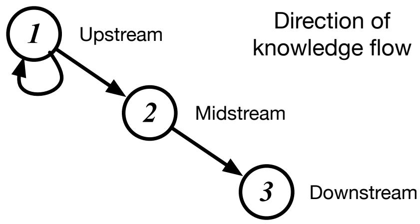
$$\mathcal{L}(\boldsymbol{b}, \xi) = \exp\left(\frac{\lambda}{\rho}\xi\boldsymbol{\gamma}'(\ln\boldsymbol{\gamma} - \ln\boldsymbol{b})\right),$$

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 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.



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

In the decentralized balanced growth path, exactly one-third of R&D resources is allocated to each sector, as in Proposition 6.

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 $\rho/\lambda = 0.1$, $\gamma_1/\gamma_3 \approx 30.1$. The difference in the knowledge stock growth rates between the efficient and the decentralized BGP is

$$g(\boldsymbol{\gamma}') - g(\boldsymbol{\beta}') = \ln\left(3 + 3\rho/\lambda + (\rho/\lambda)^2\right) - 2\ln(1+\rho/\lambda),$$

which is decreasing in ρ/λ : the economy grows faster when the planner is more patient.

B.2 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 $V\left(\boldsymbol{q}_0; \left\{\bar{\ell}_t\right\}, \{\bar{s}_t\}\right)$ denote the planner's value function, where the masses of workers and scientists are both exogenous along the entire growth path. The value function (10) 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

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

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.

B.3 Embedding Input-Output Linkages into Production Functions

In this section we introduce input-output linkages into the baseline model in Section 2 of the main text, and we show Lemma 1 and Proposition 1 extend intuitively.

Specifically, assume that producing sectoral composite good i requires using goods from other sectors as inputs:

$$\ln y_{it} = \sum_{j=1}^{K} \sigma_{ij} \ln m_{ijt} + \left(1 - \sum_{j=1}^{K} \sigma_{ij}\right) \int_0^1 \left(\ln q_{it}(\nu) + \ln x_{it}(\nu)\right) \mathrm{d}\nu, \tag{A13}$$

where intermediate varieties in sector i are produced directly from labor ($x_{it}(\nu) = \ell_{it}(\nu)$) as in the main text, m_{ijt} is the quantity of good j used for the production of good i, and σ_{ij} is the output elasticity of sector i with respect to input j. The market clearing condition of sectoral composite good follows

$$y_{jt} = \sum_i m_{ijt} + c_{jt}. \tag{A14}$$

The aggregate consumption bundle follows:

$$\ln c_t = \sum_{i=1}^{K} \delta_i \ln c_{it}. \tag{A15}$$

Because worker allocation is constant across varieties, we can rewrite (A13) as

$$\ln y_{it} = \sum_{j=1}^{K} \sigma_{ij} \ln m_{ijt} + \alpha_i \left[\ln q_{it} + \ln \ell_{it}\right], \tag{A16}$$

where $\alpha_i \equiv 1 - \sum_{j=1}^{K} \sigma_{ij}$ is the elasticity of output in sector i with respect to value-added inputs.

Substituting (A16) into (A14) and using vector notation, we have

$$\ln y_{it} = \sum_{j=1}^{K} \sigma_{ij} \ln m_{ijt} + \alpha_i \left[\ln q_{it} + \ln \ell_{it}\right].$$

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

$$\ln c^* \left(\boldsymbol{q}_t\right) \equiv \max_{\{\ell_{it}\}} \sum_{i=1}^{K} \delta_i \ln c_{it}$$

subject to (A14) and (A16). 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

$$\ln c^* \left(\boldsymbol{q}_t\right) = \text{const} + \ln \bar{\ell} + \sum_i \beta_i \ln q_{it},$$

where $\beta_i \equiv \alpha_i \left[\boldsymbol{\delta}' \left(\boldsymbol{I} - \boldsymbol{\Sigma}\right)^{-1}\right]_i$ is the product between sectoral value-added elasticity α_i and the i-th entry of the influence vector $\boldsymbol{\delta}' \left(\boldsymbol{I} - \boldsymbol{\Sigma}\right)^{-1}$. β_i can be interpreted as the value-added in sector i as a share of GDP, since the optimal worker allocation follows $\ell_{it}/\bar{\ell} = \beta_i$. Hence, results in the main text extend intuitively to this setting with input-output linkages: worker allocation follows the vector $\boldsymbol{\beta}$, and scientist allocation $\gamma_{it} \equiv s_{it}/\bar{s}$ follows $\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Omega}}{1+\rho/\lambda}\right)^{-1}$.

B.4 General Preferences and Production Functions

Our models in the main text assume specific functional forms for preferences and technologies in order to derive the optimal R&D allocations along the economy's entire growth path. In this section we relax some functional form assumptions and show the optimal R&D allocations characterized by Proposition 1 (closed economy) and Proposition 8 (open economy) still hold locally around a BGP.

B.4.1 Closed Economy

Consider a closed economy characterized by the following general preferences

$$\int_0^\infty e^{-\rho t} \mathcal{U}\left(y_t\right) \, \mathrm{d}t$$

where the consumption good is produced by a constant-returns-to-scale function $y_t = \mathcal{F}\left(\{y_{it}\}\right)$. Suppose the innovation production function is

$$n_{it} = \eta_i s_{it} \mathcal{X}_i \left(\{q_{jt}\}\right),$$

where $\mathcal{X}_i(\cdot)$ is a constant-returns-to-scale aggregator for how R&D in sector i benefits from other sectors' knowledge. The law of motion for knowledge stock in the economy is $\frac{\mathrm{d} \ln q_{it}}{\mathrm{d}t} = \lambda \ln(n_{it}/q_{it})$ and the production structure is $y_{it} = q_{it}\ell_{it}$, both the same as in the main text.

To solve the planner's problem, formulate the Hamiltonian:

$$\begin{aligned} H &= \mathcal{U}(\mathcal{F}(\{y_{it}\})) + \lambda \sum_{i=1}^{K} \mu_{it}(\ln \bar{s} + \ln \eta_i + \ln \gamma_{it} + \ln \mathcal{X}_i(\{q_{jt}\}) - \ln q_{it}) \\ &\quad + \zeta_t\left(1 - \sum_i \gamma_{it}\right) + \delta_t\left(\bar{\ell} - \sum_i \ell_{it}\right) \end{aligned}$$

By the maximum principle

$$H_{\ell_{it}} = 0 \iff \frac{\partial \mathcal{U}_t}{\partial \ln \mathcal{F}_t}\frac{\partial \ln \mathcal{F}_t}{\partial \ln y_{it}}\frac{\partial \ln y_{it}}{\partial \ln \ell_{it}} = \delta_t \ell_{it} \tag{A17}$$

$$H_{\gamma_{it}} = 0 \iff \frac{\lambda \mu_{it}}{\gamma_{it}} = \zeta_t \tag{A18}$$

$$H_{\ln q_{it}} = \rho\mu_{it} - \dot{\mu}_{it} \iff \rho\mu_{jt} - \dot{\mu}_{jt} = \frac{\partial \mathcal{U}_t}{\partial \ln \mathcal{F}_t}\frac{\partial \ln \mathcal{F}_t}{\partial \ln y_{jt}}\frac{\partial \ln y_{jt}}{\partial \ln q_{jt}} - \lambda\mu_{jt} + \lambda\sum_i \mu_i \frac{\partial \ln \mathcal{X}_{it}}{\partial \ln q_{jt}} \tag{A19}$$

Let $\beta_{it} \equiv \frac{\partial \ln \mathcal{F}_t}{\partial \ln y_{it}}$ and $\omega_{ijt} \equiv \frac{\partial \ln \mathcal{X}_{it}}{\partial \ln q_{jt}}$. First, note that (A17) and (A18) $\ell_{it} = \beta_{it}\bar{\ell}$ and $\gamma_{it} = \frac{\mu_{it}}{\sum_j \mu_{jt}}\bar{s}$. Along a BGP, $\dot{\mu}_{jt} = 0$, and both $\boldsymbol{\beta}$ and $\boldsymbol{\Omega} \equiv [\omega_{ij}]$ are time-invariant. Equation (A19) implies

$$(\rho + \lambda)\mu_j = c_t\beta_j + \lambda\sum_i \mu_i\omega_{ij}$$

for some constant c_t; hence, in vector form,

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

Because $\mathcal{X}_i(\cdot)$ has constant returns, $\boldsymbol{\Omega}$ is a row-stochastic matrix; hence

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

Note that the optimal R&D allocation along the BGP has the same form as characterized by Proposition 1, except in this more general model, $\boldsymbol{\Omega}$ is a matrix of innovation elasticities $\frac{\partial \ln \mathcal{X}_i}{\partial \ln q_j}$ evaluated locally around the BGP, and β_i is the elasticity of aggregate output with respect to the output of each sectoral good, also evaluated locally around the BGP.

B.4.2 Open Economy

Consider an open economy characterized by the following general preferences

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

Domestic production function is $y_t = \mathcal{F}\left(\{y_{it}\}\right)$, and the innovation production function is

$$n_{it} = \eta_i s_{it}\mathcal{X}_i\left(\left\{q_{jt}, q_{jt}^f\right\}\right),$$

where $\mathcal{X}_i\left(\cdot\right)$ is a constant-returns-to-scale aggregator. As in the main text, the law of motion for knowledge stock in the economy is $\frac{\mathrm{d}\ln q_{it}}{\mathrm{d}t} = \lambda\ln\left(n_{it}/q_{it}\right)$ and the sectoral production function is $y_{it} = q_{it}\ell_{it}$. The economy can import foreign goods c_t^f by exporting the unconsumed domestic goods $\left(y_t - c_t^d\right)$. Let p_t^f denote relative prices of foreign goods. Trade balance implies $p_t^f c_t^f = y_t - c_t^d$.

We look for the planner's solution along an international BGP in which $\{q_{jt}\}$ and $\left\{q_{jt}^f\right\}$ grow at the same rate, and p_t^f is time-invariant. Note that the paths of foreign knowledge and import prices $\left\{q_{jt}^f, p_t^f\right\}$ are exogenous to the planner. We formulate the Hamiltonian:

$$\begin{aligned} H &= \mathcal{U}\left(c_t^d, c_t^f\right) + \lambda\sum_{i=1}^{K}\mu_{it}\left(\ln\bar{s} + \ln\eta_i + \ln\gamma_{it} + \ln\mathcal{X}_i\left(\left\{q_{jt}, q_{jt}^f\right\}\right) - \ln q_{it}\right) \\ &\quad +\zeta_t\left(1 - \sum_i\gamma_{it}\right) + \delta_t\left(\bar{\ell} - \sum_i\ell_{it}\right) + \alpha_t\left(\mathcal{F}\left(\{y_{it}\}\right) - c_t^d - p_t^f c_t^f\right) \end{aligned}$$

By the maximum principle,

$$H_{c_t^d} = 0 \iff \frac{\partial\mathcal{U}}{\partial c_t^d} = \alpha_t$$

$$H_{c_t^f} = 0 \iff \frac{\partial\mathcal{U}}{\partial c_t^f} = p_t^f\alpha_t$$

$$H_{\ell_{it}} = 0 \iff \alpha_t\mathcal{F}_t\frac{\partial\ln\mathcal{F}_t}{\partial\ln y_{it}}\frac{\partial\ln y_{it}}{\partial\ln\ell_{it}} = \delta_t\ell_{it} \tag{A20}$$

$$H_{\gamma_{it}} = 0 \iff \frac{\lambda\mu_{it}}{\gamma_{it}} = \zeta_t \tag{A21}$$

$$H_{\ln q_{it}} = \rho\mu_{it} - \dot{\mu}_{it} \iff \rho\mu_{jt} - \dot{\mu}_{jt} = \mathcal{F}_t\frac{\partial\ln\mathcal{F}_t}{\partial\ln y_{jt}}\frac{\partial\ln y_{jt}}{\partial\ln q_{jt}} - \lambda\mu_{jt} + \lambda\sum_i\mu_i\frac{\partial\ln\mathcal{X}_{it}}{\partial\ln q_{jt}} \tag{A22}$$

Let $\beta_{it} \equiv \partial\ln\mathcal{F}/\partial\ln y_{it}$ and $\theta_{ijt} \equiv \frac{\partial\ln\mathcal{X}_{it}}{\partial\ln q_{jt}}$. Note that (A20) and (A21) $\ell_{it} = \beta_{it}\bar{\ell}$ and $\gamma_{it} = \frac{\mu_{it}}{\sum_j\mu_{jt}}\bar{s}$.

Along a BGP, $\dot{\mu}_{jt} = 0$, and both $\boldsymbol{\beta}$ and $\boldsymbol{\Theta} \equiv [\theta_{ij}]$ are time-invariant. Equation (A19) implies

$$(\rho + \lambda)\, \mu_j = c_t \beta_j + \lambda \sum_i \mu_i \theta_{ij}$$

for some constant c_t; hence, in vector form,

$$\boldsymbol{\gamma}' \propto \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1 + \rho/\lambda} \right)^{-1}. \tag{A23}$$

$\boldsymbol{\Omega}$ is a matrix of innovation elasticities with respect to domestic knowledge ($\frac{\partial \ln \mathcal{X}_i}{\partial \ln q_j}$) evaluated locally around the BGP, and β_i is the elasticity of aggregate output with respect to the output of each sectoral good, also evaluated locally around the BGP. Note that because the economy benefits from foreign spillovers, $\boldsymbol{\Theta}$ is row-substochastic, so $\xi \equiv \frac{\rho}{\rho+\lambda} \boldsymbol{\beta}' \left(\boldsymbol{I} - \frac{\boldsymbol{\Theta}}{1+\rho/\lambda} \right)^{-1} \mathbf{1} \leq 1$. When $\frac{\partial \ln \chi_{it}}{\partial \ln q_{jt}} = \omega_{ij} x_{ij}$, $\boldsymbol{\Theta} = \boldsymbol{\Omega} \circ \boldsymbol{X}$ as in Proposition 8.

B.5 Constrained Optimal Allocation in an Open Economy

Proposition 5 in Section 2.5 characterizes the constrained optimal R&D allocations when the planner can only allocate R&D resources in a subset of sectors $\mathcal{K} \subset \{1, \ldots, K\}$ and provides the formula for the welfare cost of R&D misallocation in this setting. The next result extends the analysis to an open economy.

Proposition 10. *Consider an open economy with R&D self-sufficiency measure ξ and given paths of foreign knowledge and relative import prices $\left\{ \boldsymbol{q}_t^f, p_t^f \right\}_{t=0}^{\infty}$. 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 cost of following R&D allocation $\boldsymbol{b}$ instead of the constrained optimal allocation is $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\lambda}{\rho} \xi \left(\sum_{j \in \mathcal{K}} \gamma_j \right) \left(\boldsymbol{\gamma}^{\mathcal{K}} \right)' \left(\ln \boldsymbol{\gamma}^{\mathcal{K}} - \ln \boldsymbol{b}^{\mathcal{K}} \right) \right)$.*

Proof. Following equation (A12), the welfare differences between two time-invariant paths of R&D allocation in an open economy is

$$V\left(\widetilde{\boldsymbol{b}}\right) - V(\boldsymbol{b}) = \frac{\lambda}{\rho^2} \xi \boldsymbol{\gamma}' \left(\ln \widetilde{\boldsymbol{b}} - \ln \boldsymbol{b} \right).$$

This proposition thus follows from using the same argument as in the closed economy setting; see Section A.7. □

B.6 Optimal 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

$$V = \int_0^\infty e^{-\rho t} \left(\sigma^h \ln c_t^{hh} + \left(1-\sigma^h\right) \ln c_t^{hf}\right) \,\mathrm{d}t, \tag{A24}$$

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 aggregations of intermediate varieties produced from labor (equations 2, 3, and 4). We can simplify the home production functions as

$$\ln y_t^h = \sum_i \beta_i \left(\ln q_{it}^h + \ln \ell_{it}^h\right). \tag{A25}$$

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

$$n_{it}^h = \eta_i^h s_{it}^h \chi_{it}^h, \qquad \text{where } \chi_{it}^h = \prod_{j=1}^K \left[\left(q_{jt}^h\right)^{x_{ij}^h} \left(q_{jt}^f\right)^{1-x_{ij}^h}\right]^{\omega_{ij}}, \tag{A26}$$

and the law of motion for home knowledge stock is

$$\frac{\mathrm{d} \ln q_{it}^h}{\mathrm{d}t} = \lambda \ln \left(n_{it}^h / q_{it}^h\right). \tag{A27}$$

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 (A24), we can write the home planning problem as

$$V^* \left(\left\{ \boldsymbol{\ell}_t^f, \boldsymbol{s}_t^f \right\}_{t=0}^{\infty} \right) \equiv \max_{\left\{ s_{it}^h, \ell_{it}^h \right\}} \int_0^\infty e^{-\rho t} \left(\sigma^h \ln y_t^h + \left(1-\sigma^h\right) \ln y_t^f \right) \mathrm{d}t, \tag{A28}$$

subject to the innovation production functions (A27 and A26), goods production function (A25), 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 consumer expenditure shares captured by

$$\hat{\boldsymbol{\beta}} \equiv \left[\sigma^h \boldsymbol{\beta}', \left(1-\sigma^h\right) \boldsymbol{\beta}' \right], \tag{A29}$$

and the innovation network captured by

$$\hat{\boldsymbol{\Omega}} \equiv \begin{bmatrix} \boldsymbol{\Omega} \circ \boldsymbol{X}^h & \boldsymbol{\Omega} - \boldsymbol{\Omega} \circ \boldsymbol{X}^h \\ \boldsymbol{\Omega} - \boldsymbol{\Omega} \circ \boldsymbol{X}^f & \boldsymbol{\Omega} \circ \boldsymbol{X}^f \end{bmatrix}. \tag{A30}$$

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

$$\hat{\gamma}' \equiv \frac{\rho}{\rho + \lambda} \left(\boldsymbol{I}_{2K \times 2K} - \frac{\hat{\boldsymbol{\Omega}}}{1 + \rho/\lambda} \right)^{-1}. \tag{A31}$$

Next, recognize that the actual home planner's problem (A28) is essentially the same as in the hypothetical integrated economy, but with the additional constraint that the home planner can only allocate resources domestically. Hence, our results in Proposition 5 directly apply. We summarize the result into 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 (A28) 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 cost of following R&D allocation $\boldsymbol{b}$ *instead of the optimal allocation is* $\mathcal{L}^{\mathcal{K}}(\boldsymbol{b}) = \exp\left(\frac{\lambda}{\rho}\left(\sum_{j\in\mathcal{K}} \hat{\gamma}_j\right)\left(\hat{\boldsymbol{\gamma}}^{\mathcal{K}}\right)'\left(\ln \hat{\boldsymbol{\gamma}}^{\mathcal{K}} - \ln \boldsymbol{b}\right)\right)$.

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. Innovation Data

U.S. patent data are obtained from the United States Patent and Trademark Office (USPTO).[12] 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 Innovation 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

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

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),[13] we use the country of 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 Innovation 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.

[13] 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%).

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.[14] 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.[15] For each sector, we can aggregate all 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.

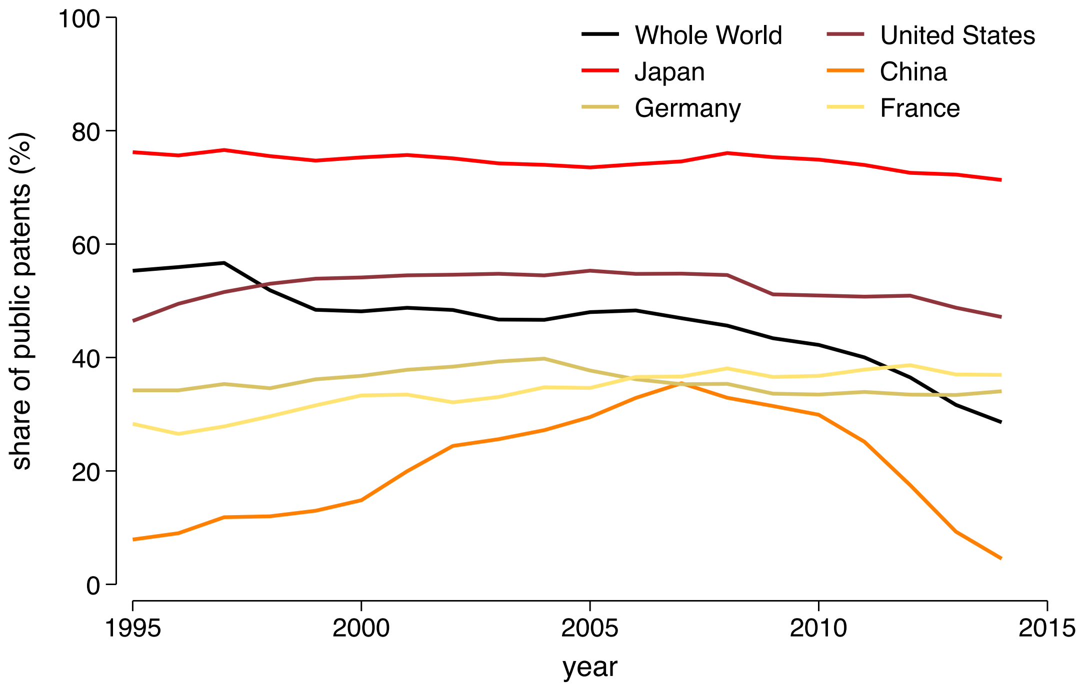
The benefit of using information on public firms to accurately link innovation to industrial sectors warrants the question of how representative those firms' innovation are. We find that firms in our data set produce about half of all patents in each country—for example, our sample of firms covers 44% 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 public firms in the whole world and in different countries. The similarity of industry distribution between patents from public 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 public firms.

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

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

[15] 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 public firms across the world and in different countries.

(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.

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).

- **Publication (most basic unit):** After an application is filed, various publications could be issued.[16] 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:**[17] 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.2 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

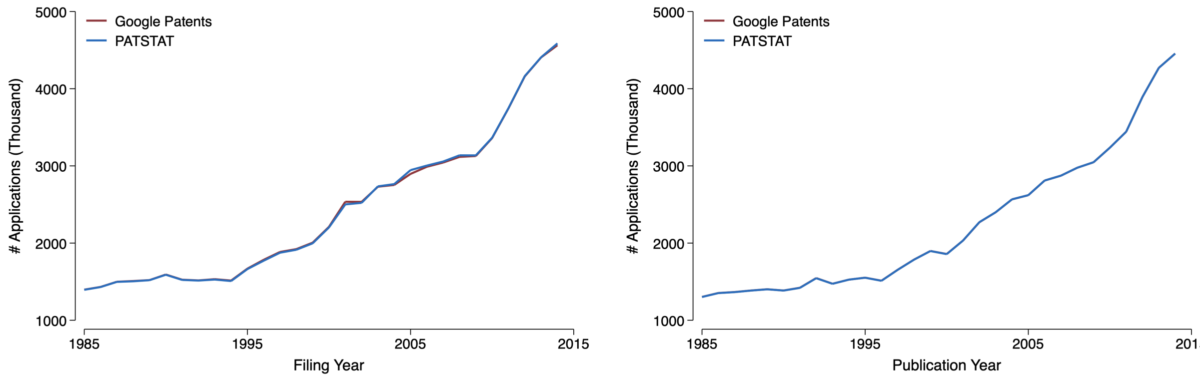
Publications represent the most comprehensive set of innovation-related documents, yet many of them are irrelevant for studying innovation—some publications are associated with denied applications, some are design patents unrelated to scientific or technological progress, etc. As a result, it is useful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

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

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

Figure A.2. Google Patents v.s. PATSTAT by Year

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



in the USPTO 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.[18] 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.1 shows the comparison of granted patents between Google Patents and PATSTAT and list the sources of coverage differences.

[18]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.1. 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

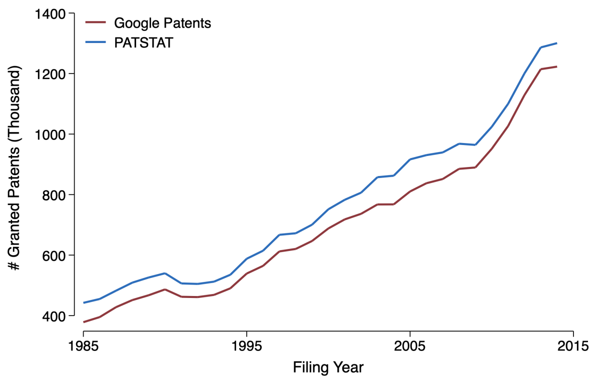
	# Patents			
Patents granted in 1985—2014	18,144,529	100.00%		
Overlapped with PATSTAT	17,135,612	94.44%		
Non-overlapped with PATSTAT	1,008,917	5.56%	100.00%	
1. Additional patent office data from legal event data	0		0.00%	
2. Additional rules used to identify granted patents	1,008,917		100.00%	100.00%
(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.

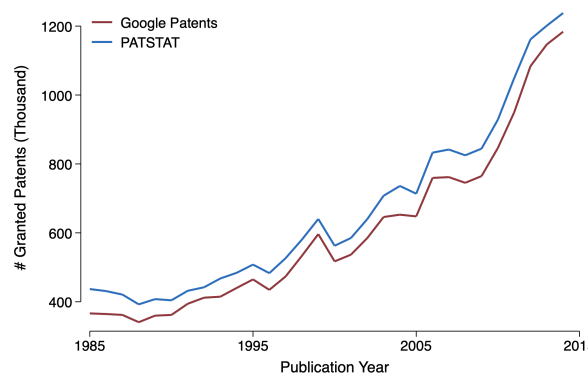
Despite those differences, Google Patents and PATSTAT agree on roughly 95% of the identified granted patents. In Figure A.3, we present the numbers of granted patents in Google Patents and PATSTAT. We also show this difference across various patent offices and countries of origin.

Figure A.3. Google Patents v.s. PATSTAT Coverage

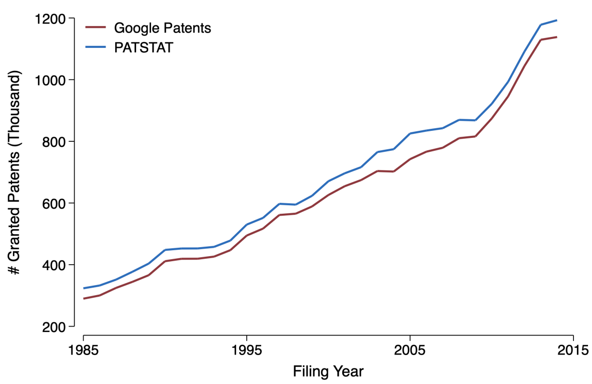
(a) Granted Patents by Filing Year



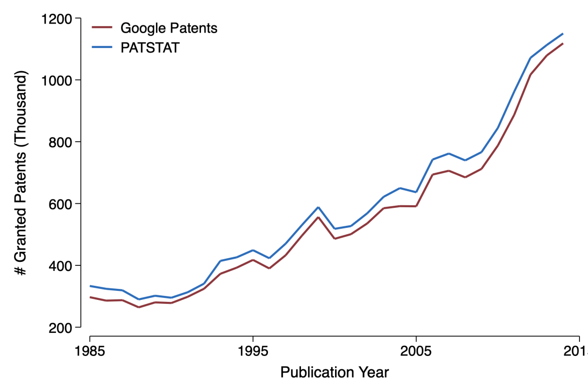
(b) Granted Patents by Publication Year



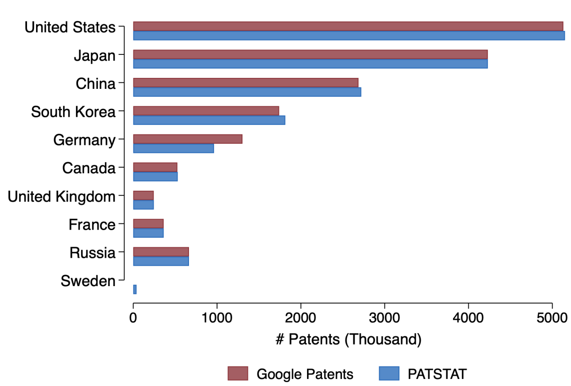
(c) Granted Patents by Filing Year (43 WIOD Countries)



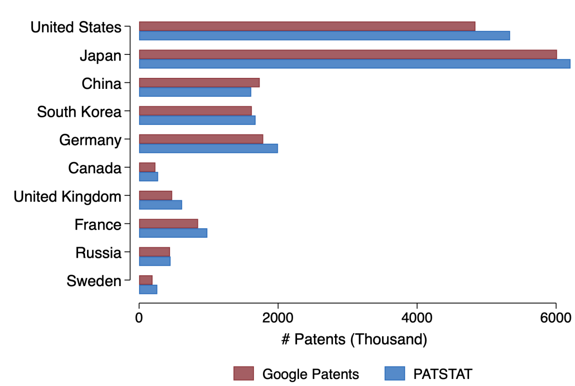
(d) Granted Patents by Publication Year (43 WIOD Countries)



(e) Google Patents vs. PATSTAT By Patent Offices



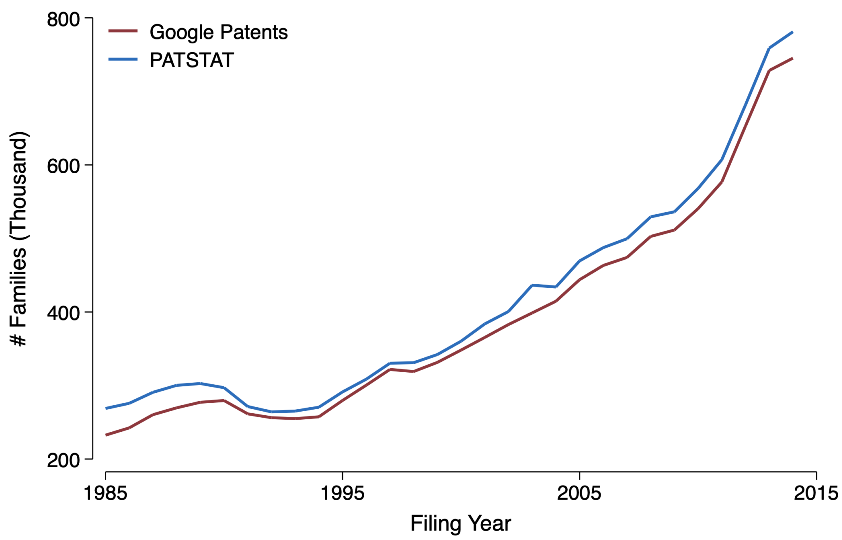
(f) Google Patents vs. PATSTAT By Invention Origin



D.3 Patent Family

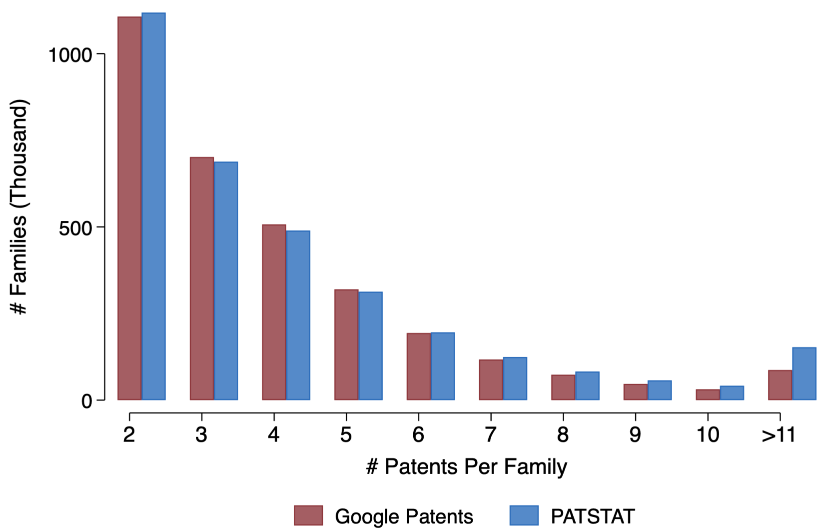
Defining patent family involves the use of information regarding priority dates and priority patents in the global patent database, among others. Figure A.4 presents the number of patent families identified in both data sets. They are very comparable to each other, and the minor gap can be explained by the differences in the number of identified patents described in the previous section.

Figure A.4. Google Patents v.s. PATSTAT: Patent Families by Year



To further check this consistency, in Figure A.5 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.5. 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.2), and they have virtually identical properties such as centrality (Figure A.6) and visualizations (Figure A.7).

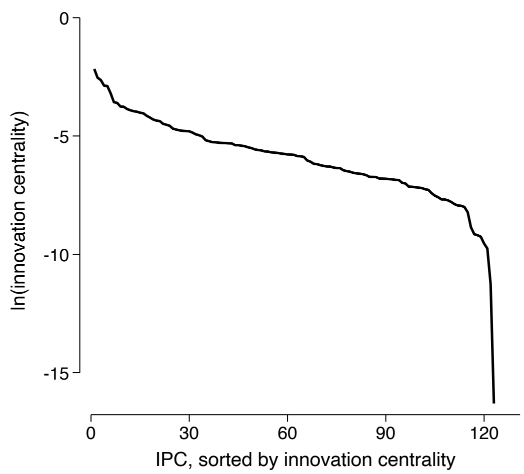
Table A.2. Correlations of Between the Innovation Network from Google Patents and PATSTAT

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.6. Innovation Centrality and Key Sectors for PASTAT

(a) Innovation Centrality Across IPCs



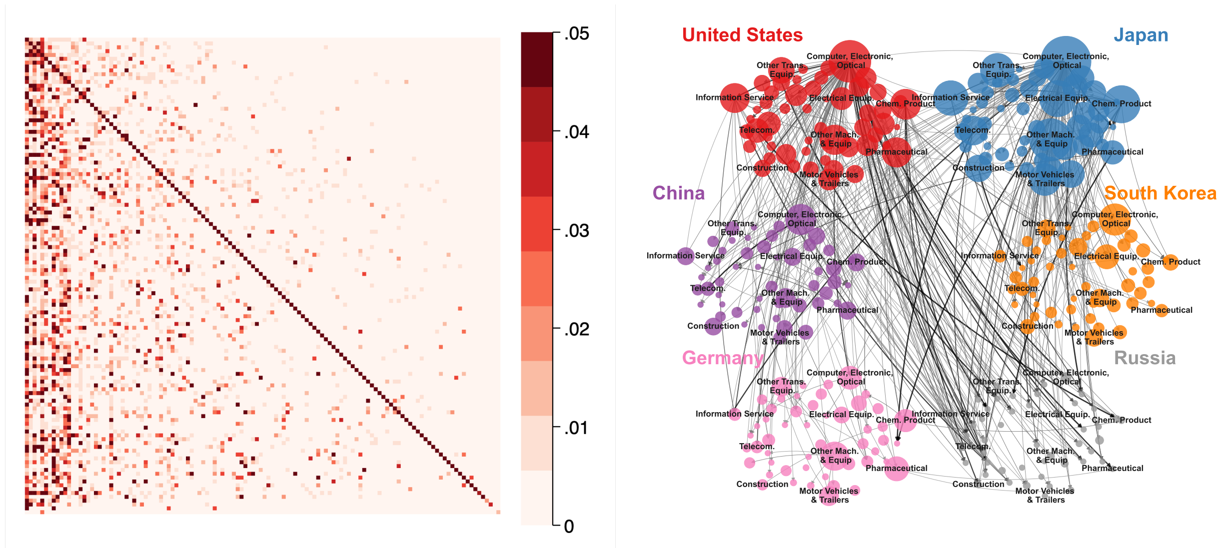
(b) Top Ten IPCs by Innovation Centrality a_i

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.7. 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 $\mathbf{\Omega}$ 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.3. Evidence of the Global Innovation Network for Knowledge Spillovers
Based on WIOD - PATSTAT

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge^{Up}_{mit}$	0.181***	0.193***	0.174***	0.285***	0.325***	0.275***
	(0.053)	(0.056)	(0.054)	(0.080)	(0.082)	(0.081)
$ln(R\&D)_{mi,t-1}$	0.031***	0.031***	0.030***	0.036***	0.038***	0.036**
	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
$Knowledge^{Down}_{mit}$		-0.035			-0.113***	
		(0.030)			(0.039)	
$Knowledge^{Up,IO}_{mit}$			0.054			-0.036
			(0.067)			(0.071)
R^2	0.968	0.968	0.969	0.943	0.943	0.944
No. of Country x Sectors	564	564	550	564	564	550
No. of Obs	10549	10549	10315	10549	10549	10315
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, $\text{Knowledge}_{it}^{Up}$, is the knowledge from upstream, defined in (26). 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 (23) 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.4 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.4 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.4. 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 (23) but restrict the sample to all patents from each country. Table A.5 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.5. 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 5.2 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.6), additional measures of innovation output (Table A.7), and different time horizons to calculate upstream innovation (Tables A.9 and A.10).

Table A.6. 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)
$Knowledge_{it}^{Up}$	0.519***	0.548***	0.050***	0.075***
	(0.083)	(0.106)	(0.011)	(0.015)
$ln(R\&D)_{i,t-1}$	0.298***	0.254**	0.006*	0.004
	(0.075)	(0.113)	(0.003)	(0.005)
$Knowledge_{it}^{Down}$	-0.243***	-0.348***	-0.032***	-0.025**
	(0.064)	(0.090)	(0.008)	(0.011)
R^2	0.959	0.947	0.947	0.905
No. of Sectors	431	431		
No. of Country x Sectors			4,560	4,560
No. of Obs	8,620	8,620	8,2977	8,2977
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.7. U.S. Evidence of Knowledge Spillover Through Innovation Networks
Additional Innovation Measure

$Y=$	ln(Patent Value)			
	(1)	(2)	(3)	(4)
$Knowledge_{it}^{Up}$	0.929***	0.926***	0.914***	1.545***
	(0.316)	(0.329)	(0.316)	(0.447)
$ln(R\&D)_{i,t-1}$	0.307***	0.307***	0.307***	0.332***
	(0.100)	(0.099)	(0.100)	(0.124)
$Knowledge_{it}^{Down}$		0.008		
		(0.117)		
$Knowledge_{it}^{Up,IO}$			0.073	
			(0.223)	
Specification	OLS	OLS	OLS	IV 2nd Stage
R^2	0.885	0.885	0.885	0.152
No. of Sectors	94	94	94	94
No. of Obs	1,847	1,847	1,847	1,113
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.8. 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)
$Knowledge_{it}^{Up}$	0.508***		0.743***		0.914***	
	(0.174)		(0.196)		(0.316)	
$ln(R\&D)_{i,t-1}$	0.279***	0.304***	0.261***	0.297***	0.307***	0.351***
	(0.060)	(0.057)	(0.086)	(0.083)	(0.100)	(0.098)
$Knowledge_{it}^{Up,IO}$	0.363**	0.462***	0.268	0.413*	0.073	0.251
	(0.173)	(0.174)	(0.205)	(0.209)	(0.223)	(0.238)
R^2	0.917	0.914	0.902	0.898	0.885	0.880
No. of Sectors	94	94	94	94	94	94
No. of Obs	1,847	1,847	1,847	1,847	1,847	1,847
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).

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)
$Knowledge_{it}^{Up,\tau=5}$	0.454***	0.471***	0.392***	0.699***	0.730***	0.653***
	(0.152)	(0.176)	(0.149)	(0.163)	(0.173)	(0.164)
$ln(R\&D)_{i,t-1}$	0.288***	0.288***	0.291***	0.273***	0.273***	0.275***
	(0.062)	(0.061)	(0.058)	(0.084)	(0.084)	(0.084)
$Knowledge_{it}^{Down,\tau=5}$		-0.038			-0.068	
		(0.157)			(0.095)	
$Knowledge_{it}^{Up,IO}$			0.392**			0.296
			(0.172)			(0.205)
R^2	0.914	0.914	0.916	0.901	0.901	0.902
No. of Sectors	94	94	94	94	94	94
No. of Obs	1,847	1,847	1,847	1,847	1,847	1847
Fixed Effects		Sector, Year			Sector, Year	

Panel (b): $\tau = 20$

Y=	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge_{it}^{Up,\tau=20}$	0.697***	0.714***	0.611***	0.871***	0.903***	0.808***
	(0.183)	(0.206)	(0.179)	(0.209)	(0.223)	(0.200)
$ln(R\&D)_{i,t-1}$	0.270***	0.269***	0.274***	0.254***	0.253***	0.258***
	(0.062)	(0.061)	(0.059)	(0.086)	(0.085)	(0.086)
$Knowledge_{it}^{Down,\tau=20}$		-0.038			-0.071	
		(0.164)			(0.101)	
$Knowledge_{it}^{Up,IO}$			0.338*			0.249
			(0.173)			(0.204)
R^2	0.916	0.916	0.917	0.901	0.901	0.902
No. of Sectors	94	94	94	94	94	94
No. of Obs	1,847	1,847	1,847	1,847	1,847	1847
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. 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)
$Knowledge_{mit}^{Up,\tau=5}$	0.105**	0.125**	0.099*	0.227***	0.280***	0.224***
	(0.050)	(0.055)	(0.051)	(0.078)	(0.087)	(0.080)
$ln(R\&D)_{mi,t-1}$	0.034***	0.034***	0.032***	0.047***	0.048***	0.047***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.014)
$Knowledge_{mit}^{Down,\tau=5}$		-0.031			-0.085*	
		(0.036)			(0.048)	
$Knowledge_{mit}^{Up,IO}$			0.080			-0.037
			(0.064)			(0.071)
R^2	0.969	0.969	0.969	0.944	0.944	0.944
No. of Country x Sectors	564	564	550	564	564	550
No. of Obs	10,552	10,552	10,318	10,552	10,552	10,318
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)
$Knowledge_{mit}^{Up,\tau=20}$	0.199***	0.249***	0.198***	0.426***	0.499***	0.426***
	(0.059)	(0.064)	(0.059)	(0.085)	(0.092)	(0.087)
$ln(R\&D)_{mi,t-1}$	0.033***	0.034***	0.032***	0.045***	0.047***	0.045***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.013)	(0.014)
$Knowledge_{mit}^{Down,\tau=20}$		-0.138***			-0.202***	
		(0.048)			(0.066)	
$Knowledge_{mit}^{Up,IO}$			0.081			-0.034
			(0.064)			(0.070)
R^2	0.969	0.969	0.969	0.944	0.945	0.945
No. of Country x Sectors	564	564	550	564	564	550
No. of Obs	10,552	10,552	10,318	10,552	10,552	10,318
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 $\tau = 10$.

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 (25) 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.11 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.11. Predicting Sectoral Patent Count Using R&D Tax Credits

	United States	Global
$Y=$	ln(Patents) (1)	ln(Patents) (2)
$ln(User\ Cost\ of\ R\&D\ Capital)$	-11.774*** (4.041)	-0.288** (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 (24), 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 (25). We report the first-stage regressions in Table A.12, and domestic and global versions of the knowledge diffusion results in Tables A.13 and A.14, corresponding to Tables 2 and 3 in the paper.

Table A.12. Predicting Sectoral Patent Count Using R&D Tax Credits

	United States	Global
$Y=$	$Knowledge_{it}^{Up}$ (1)	$Knowledge_{mit}^{Up}$ (2)
$Knowledge_{it}^{Up,IV}$	1.110***	
	(0.051)	
$Knowledge_{mit}^{Up,IV}$		0.540***
		(0.045)
$ln(R\&D)_{i,t-1}$	0.034***	0.001
	(0.012)	(0.004)
Fixed Effects		
Sector	Yes	
Year	Yes	
Country x Sector		Yes
Country x Year		Yes
Sector x Year		Yes
F-statistics	465.9	146.2
R^2	0.983	0.982
No. of Sectors	94	
No. of Country x Sectors		280
No. of Obs	1,113	4,467

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

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge^{Up}_{it}$	0.679**	0.683**	0.687**	0.974***	0.995***	0.980***
	(0.266)	(0.268)	(0.263)	(0.279)	(0.281)	(0.277)
$ln(R\&D)_{i,t-1}$	0.269***	0.271***	0.257***	0.174**	0.184**	0.163*
	(0.070)	(0.074)	(0.068)	(0.082)	(0.086)	(0.084)
$Knowledge^{Down}_{it}$		-0.013			-0.074	
		(0.131)			(0.109)	
$Knowledge^{Up,IO}_{it}$			0.337			0.319
			(0.393)			(0.439)
R^2	0.152	0.151	0.152	0.099	0.100	0.094
No. of Sectors	94	94	94	94	94	94
No. of Obs	1,113	1,113	1,113	1,113	1,113	1,113
Fixed Effects	Sector, Year			Sector, Year		

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

Table A.14. Global Evidence of Knowledge Spillovers Through Innovation Networks–*Second-Stage IV Results*

$Y=$	ln(Patents)			ln(Cites)		
	(1)	(2)	(3)	(4)	(5)	(6)
$Knowledge^{Up}_{mit}$	0.202*	0.206*	0.222*	0.405***	0.413***	0.421***
	(0.113)	(0.109)	(0.116)	(0.147)	(0.147)	(0.151)
$ln(R\&D)_{mi,t-1}$	0.066***	0.066***	0.066***	0.072***	0.072***	0.072***
	(0.014)	(0.014)	(0.014)	(0.022)	(0.022)	(0.022)
$Knowledge^{Down}_{mit}$		-0.008			-0.018	
		(0.084)			(0.118)	
$Knowledge^{Up,IO}_{mit}$			-0.130			-0.057
			(0.380)			(0.378)
R^2	0.040	0.040	0.032	0.031	0.032	0.029
No. of Country x Sectors	280	280	275	280	280	275
No. of Obs	4,467	4,467	4,412	4,467	4,467	4,412
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 6.

- Tables A.15 and A.16 present cross-country and time-series correlations of optimal R&D allocation γ.
- Figures A.8 and A.12 present analysis using patent outputs as innovation allocation measures, supplementing analysis using R&D expenditure shares in the paper.
- Figure A.10 presents analysis using alternative parameters of ρ/λ; Figure A.9 presents analysis using data from different years.
- Figure A.11 and Table A.17 present welfare calculation using all sectors in an economy.
- Figure A.13 provides additional analysis on the time series of R&D misallocation and implied welfare cost.

Table A.15. Unilaterally Optimal R&D Allocations Across Countries

Countries	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
US		0.98	0.91	0.94	0.94	0.85	0.92	0.92	0.93	0.89
Japan	0.89		0.94	0.96	0.93	0.87	0.89	0.90	0.90	0.86
China	0.89	0.93		0.96	0.86	0.90	0.82	0.84	0.83	0.76
South Korea	0.87	0.91	0.93		0.86	0.82	0.81	0.83	0.82	0.77
Germany	0.77	0.91	0.80	0.74		0.83	0.98	0.97	0.93	0.93
Russia	0.85	0.78	0.84	0.71	0.72		0.80	0.82	0.85	0.74
France	0.82	0.78	0.78	0.66	0.81	0.81		0.98	0.93	0.95
UK	0.69	0.69	0.68	0.58	0.79	0.65	0.94		0.93	0.94
Canada	0.91	0.81	0.81	0.70	0.78	0.89	0.89	0.78		0.89
Netherlands	0.49	0.52	0.51	0.43	0.61	0.40	0.80	0.92	0.62	

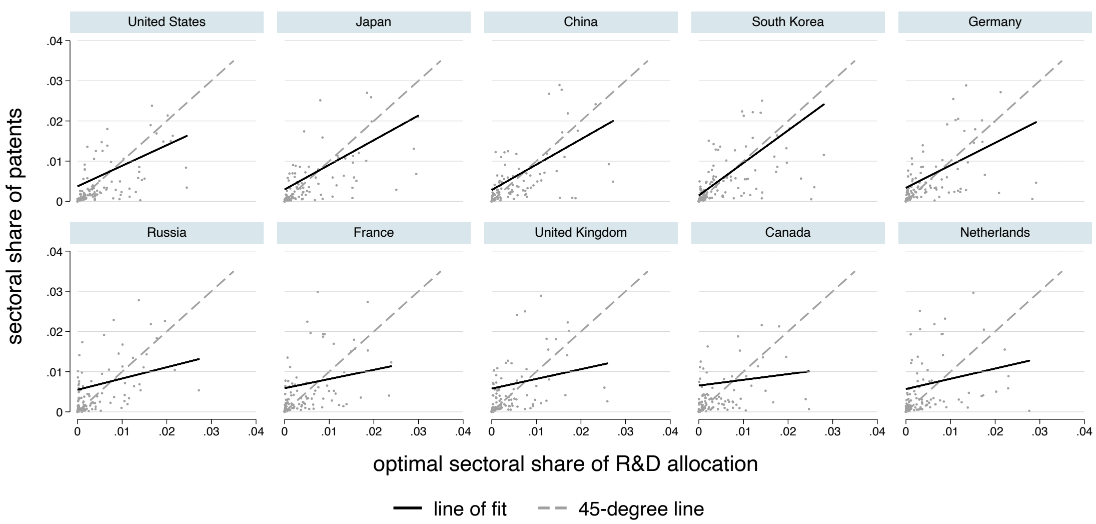
Notes. This table shows the pair-wise correlations of optimal R&D allocations γ across countries using country-specific statistics as of 2010. The lower triangular panel shows the Pearson correlation coefficients; the upper triangular panel shows Spearman's rank correlation.

Table A.16. Unilaterally Optimal US R&D Allocations of Across Time

Time Period	2020	2010	2000	1990	1980
2020		1.00	0.99	0.99	0.98
2010	0.99		0.99	0.98	0.98
2000	0.92	0.94		1.00	1.00
1990	0.92	0.93	1.00		1.00
1980	0.91	0.92	1.00	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.

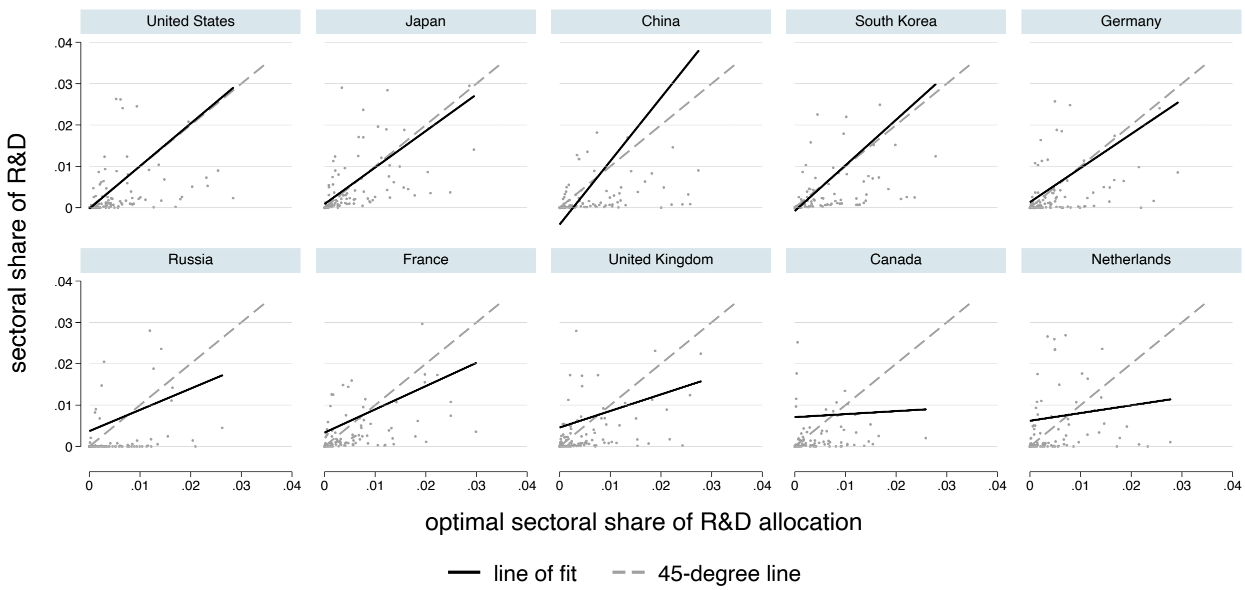
Figure A.8. Alignment Between Real Allocation and Optimal Allocation Across Countries
Using Sectoral Share of Patents as Real Allocation



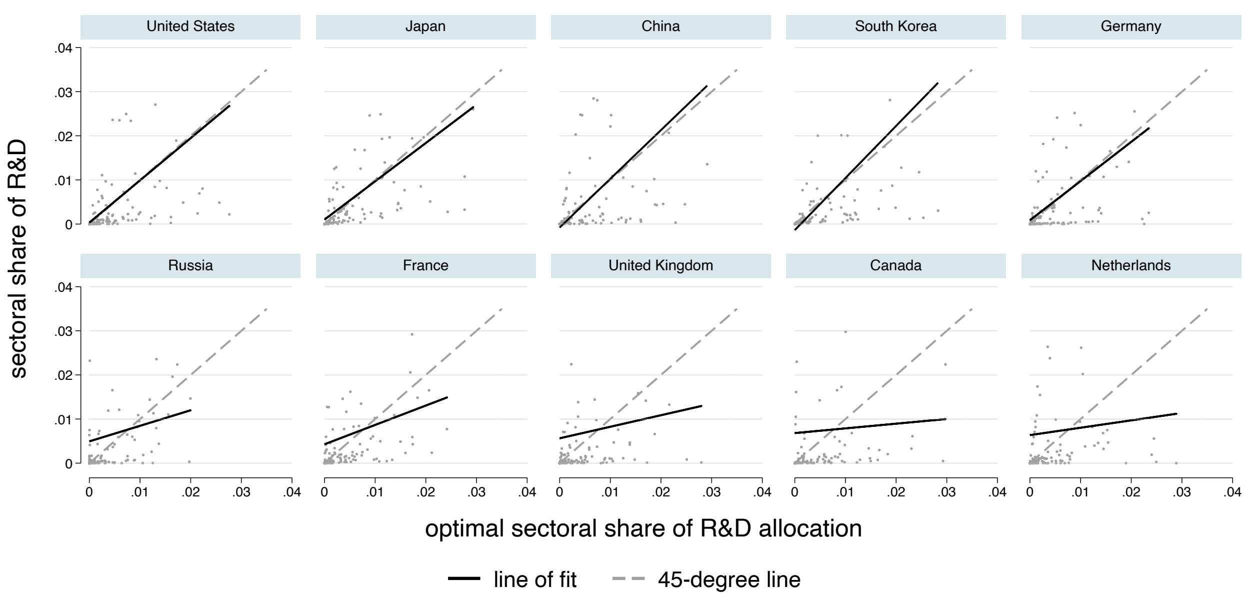
Notes. This figure reproduces Figure 7 in the paper. The figure shows scatter plots of sectoral patent output share in total patent output in the country against the optimal sectoral share of R&D allocation for top ten innovative countries in 2010. The solid line is the linear fit; the dashed line is the 45-degree line.

Figure A.9. Alignment Between Real Allocation and Optimal Allocation Across Countries
Different Years

Panel (a): 2000



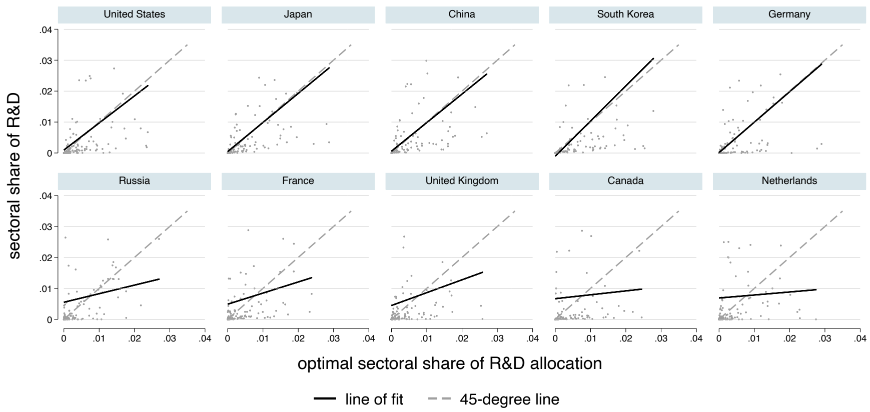
Panel (b): 2005



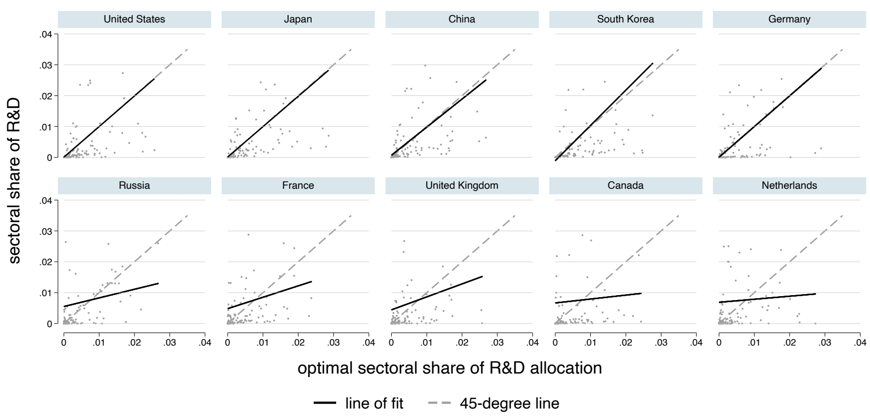
Notes. This figure reproduces Figure 7 in the paper using data from alternative years. The figure shows scatter plots of sectoral R&D expenditure share in total national R&D expenditures against the optimal sectoral share of R&D allocation for top ten innovative countries. The solid line is the linear fit; the dashed line is the 45-degree line.

Figure A.10. Alignment Between Real Allocation and Optimal Allocation Across Countries
Using Alternative Parameter Values

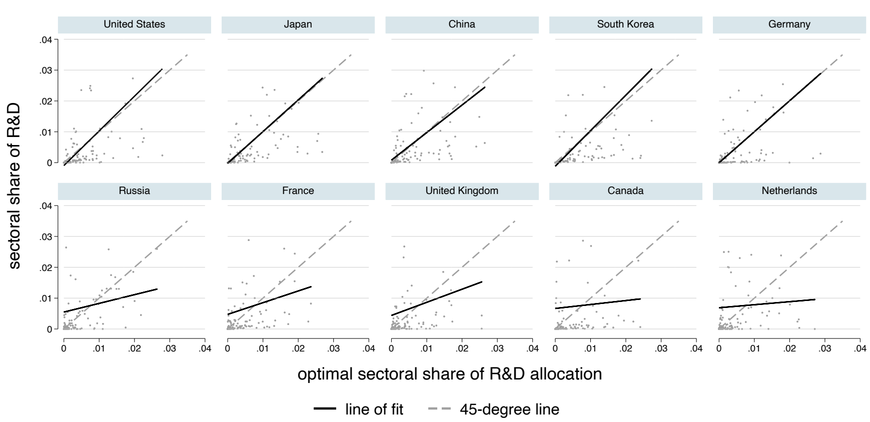
Panel (a): Use $(1+\rho/\lambda)^{-1} = 0.6$



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

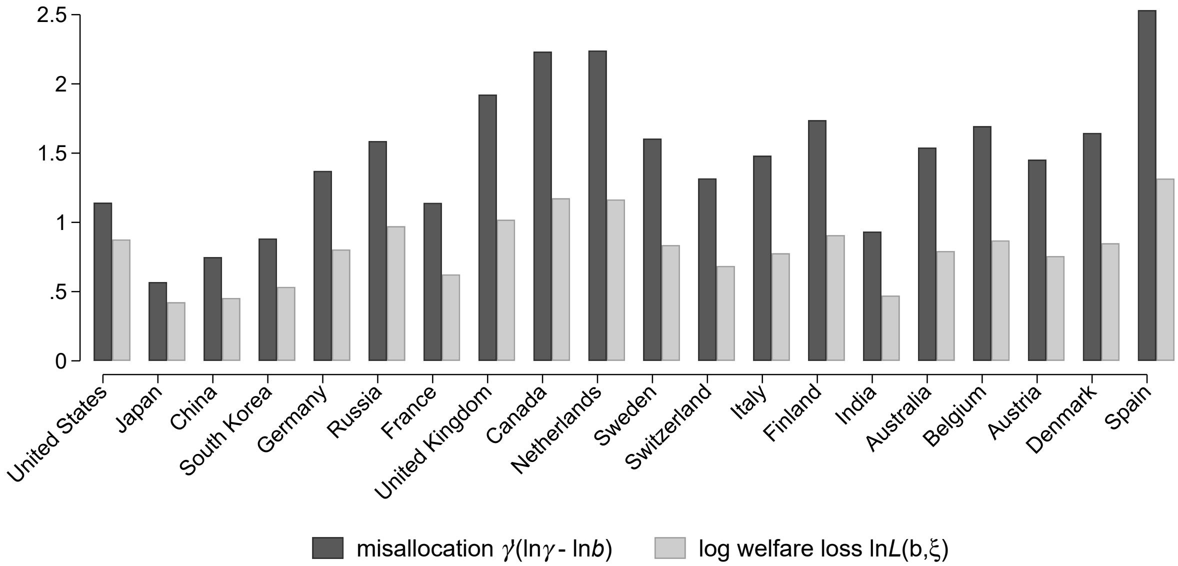


Panel (c): Use $(1+\rho/\lambda)^{-1} = 0.8$



Notes. This table reproduces Figure 7 in the paper with alternative parameter values of ρ/λ.

Figure A.11. R&D Misallocation and Welfare Cost Across Countries



Notes. This figure reproduces Figure 8 in the paper using all IPCs (instead of only top 50 ones). It shows the level of R&D misallocation (dark bars) and associated welfare cost (light grey bars) across 20 innovative countries with the highest patent outputs in our sample, using 2010-2014 data.

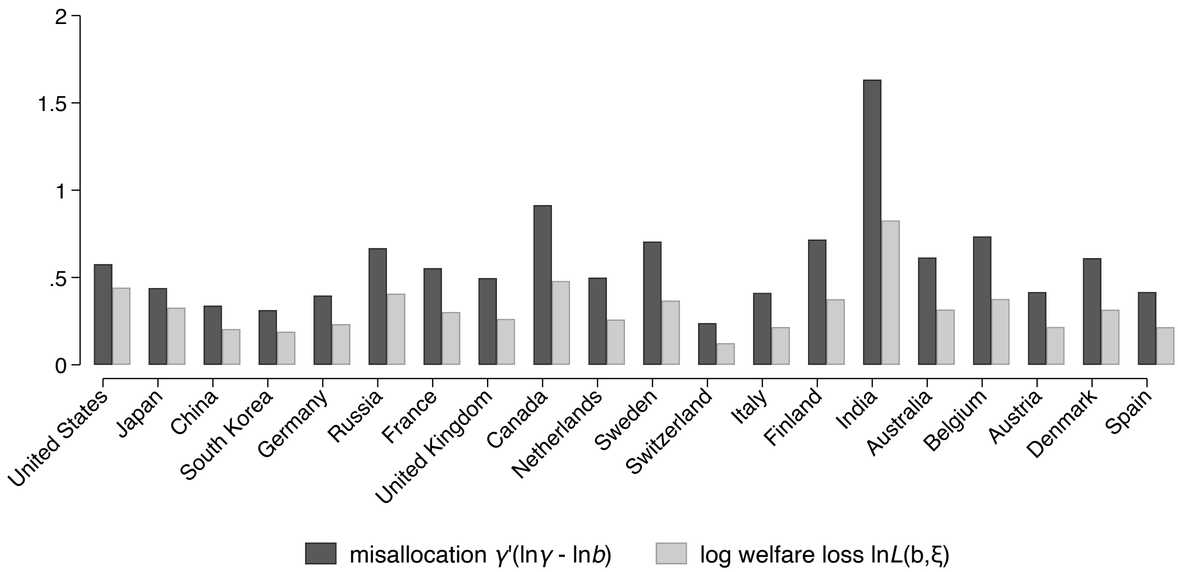
Table A.17. Percentage (%) Consumption Gains By Moving to Japan's Level of R&D Allocative Efficiency

	US	Japan	China	South Korea	Germany	Russia	France	UK	Canada	Netherlands
2000	49.31	-	149.66	54.59	63.00	78.94	47.81	72.85	332.06	81.02
2005	49.31	-	85.54	20.64	73.10	120.51	33.48	112.50	150.04	126.69
2010	55.26	-	11.52	20.97	60.15	86.57	36.78	104.90	139.84	138.66

	Sweden	Switzerland	Italy	Finland	India	Australia	Belgium	Austria	Denmark	Spain
2000	93.53	28.50	62.90	40.90	23.36	35.47	191.03	11.20	93.69	38.26
2005	59.20	44.87	114.28	50.79	22.96	45.45	124.41	31.48	79.94	34.33
2010	71.54	47.62	61.33	84.16	20.22	64.90	78.24	58.46	74.29	177.54

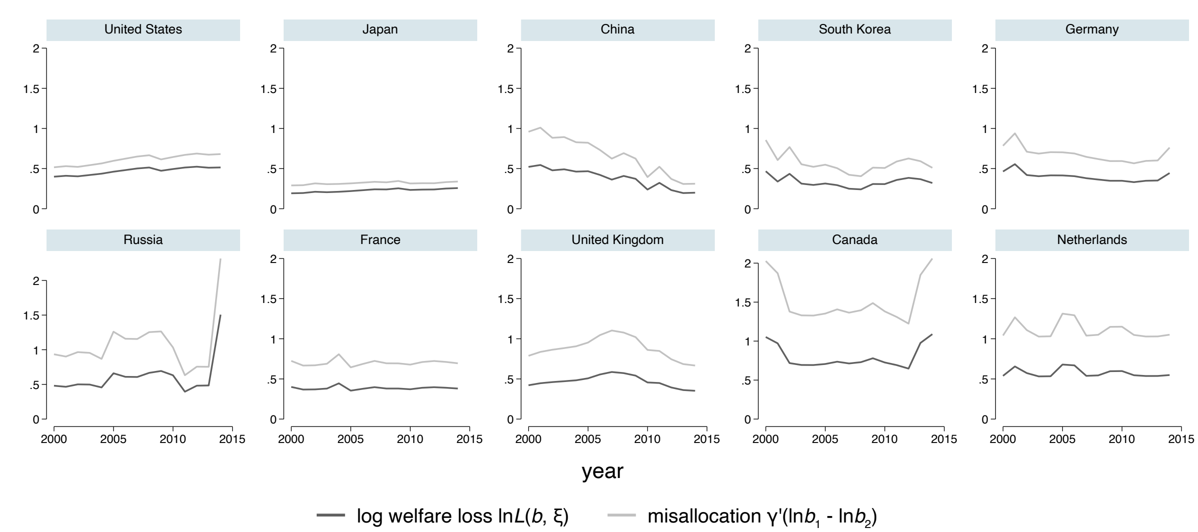
Notes. This table reproduces Table 4 in the paper using all IPCs (instead of only top 50 ones). It shows the consumption-equivalent welfare gains when each economy moves to Japan's level of R&D allocative efficiency. Specifically, this table shows the welfare gain of moving from country-specific levels of misallocation, captured by $\boldsymbol{\gamma}' (\ln \boldsymbol{\gamma} - \ln \boldsymbol{b})$, to Japan's level $\boldsymbol{\gamma}'_{JP} (\ln \boldsymbol{\gamma}_{JP} - \ln \boldsymbol{b}_{JP})$ in the corresponding year.

Figure A.12. 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 8 in the paper, but uses sectoral share of patents, rather than R&D expenditure shares, as the real allocation.

Figure A.13. 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.