NBER WORKING PAPER SERIES

BACK TO BASICS: BASIC RESEARCH SPILLOVERS, INNOVATION POLICY AND GROWTH

Ufuk Akcigit
Douglas Hanley
Nicolas Serrano-Velarde

Working Paper 19473
http://www.nber.org/papers/w19473

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2013

We thank Daron Acemoglu, Steve Bond, Giammario Impullitti, Dirk Krueger, Matthew Mitchell, Pietro Peretto, John Seater, Holger Sieg and seminar participants at AEA Meetings in Chicago, Canadian Macroeconomics Study Group, Centre for Business Taxation Symposium (Oxford), IIO Conference, Bank of Italy, Industry and Labor Market Dynamics (Barcelona), NBER Productivity Lunch Seminar, SKEMA Workshop on Economic Growth, Structural Approaches to Productivity and Industrial Dynamics (Rome), Turkish Central Bank, UCLA, University of Pennsylvania Economics, Wharton, and World Bank for their insights. We thank Claire Lelarge for very helpful comments and for providing us with patent data. Akcigit gratefully acknowledges the NBER Innovation Policy and the Economy Research Grant. Serrano-Velarde gratefully acknowledges financial support from the ESRC (Grant No RES-060-25-0033). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Back to Basics: Basic Research Spillovers, Innovation Policy and Growth
Ufuk Akcigit, Douglas Hanley, and Nicolas Serrano-Velarde
NBER Working Paper No. 19473
September 2013
JEL No. L78,O31,O38,O40,O43,O47

ABSTRACT

This paper introduces a model of endogenous growth through basic and applied research. Basic research differs from applied research in the nature and the magnitude of the generated spillovers. We propose a novel way of empirically identifying these spillovers and embed them in a general equilibrium framework with private firms and a public research sector. After characterizing the equilibrium, we estimate our model using micro-level data on research expenditures by French firms. Our key finding is that standard R&D policies can accentuate the dynamic misallocation in the economy. We also find a strong complementarity between the property rights of basic research and the optimal funding of public research.

Ufuk Akcigit
Department of Economics
University of Pennsylvania
3718 Locust Walk, #445
Philadelphia, PA 19104
and NBER
uakcigit@econ.upenn.edu

Douglas Hanley
Department of Economics,
University of Pennsylvania
3718 Locust Walk, #160
Philadelphia, PA 19104
dohan@sas.upenn.edu

Nicolas Serrano-Velarde
Bocconi University
Via Roentgen 1
20135 Milan
Italy
nicolas.serranovelarde@unibocconi.it

1 Introduction

Fostering economic growth is one of the primary objectives of economists and policymakers. The amount of resources invested in research is often at the heart of the debate regarding how to best achieve this. It is evident that the level of research investment plays an important role in the pace of long-term technological progress and economic growth, and countries allocate a significant share of their GDP to researching new products and technologies in this spirit (see Figure 1). Less well known, however, is what role the composition of this research plays in determining growth, particularly when considering the breakdown between basic and applied research.

The distinction between these two types of research is conceptually important. According to the NSF, basic research investment refers to a "systematic study to gain more comprehensive knowledge or understanding of the subject under study without specific applications in mind." Conversely, applied research is defined as a "systematic study to gain knowledge or understanding to meet a specific, recognized need."[1] This distinction is of empirical importance since almost half of total research investment is allocated to basic research in countries such as France and the US (see Figure 2).

Figure 1: Total Research to GDP Ratio in France and the US

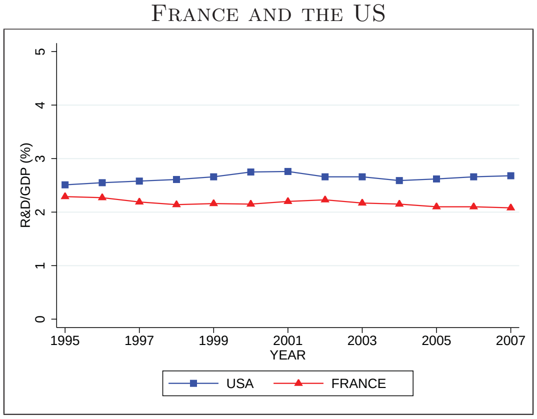
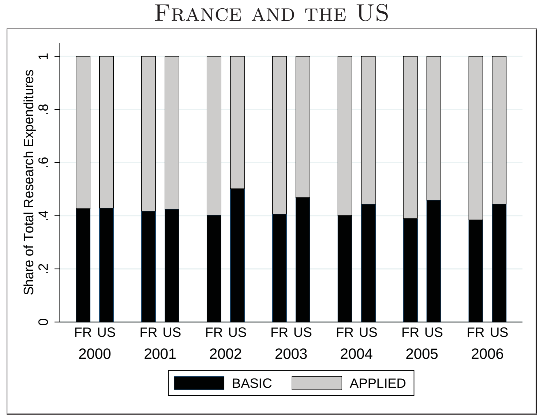


Figure 2: Composition of Research Investment in France and the US



The issue of investment in basic research also received fresh policy interest in a recent report by the US Congress's Joint Economic Committee, where it is argued that despite its value to society as a whole, basic research is underfunded by private firms precisely because it is performed with no specific commercial applications in mind. The report states that the level of federal funding for basic research is "worrisome" and should be increased (JEC (2010)).

Despite clear empirical importance and considerable policy interest, the differential nature of the roles played by basic and applied research in the growth process is still relatively unexplored,

[1] Although basic research may not have specific applications as its goal, it can be directed to fields of current or potential interest. This focus is often the case when performed by industry or mission-driven federal agencies. In industry, applied research includes investigations to discover new scientific knowledge that has specific commercial objectives with respect to products, processes, or services. http://www.nsf.gov/statistics/seind10/c4/c4s.htm#sb2

and many related questions remain to be answered: What are the key roles of basic and applied research for productivity growth? What are the incentives of private firms to do basic research? How does public basic research contribute to innovation and productivity growth? How sizable are the spillovers from basic research? What are the potential inefficiencies in a competitive economy, and what are the appropriate government policies to mitigate them and promote economic growth? Our goal in this paper is to shed light on these important questions, which have long been at the heart of macroeconomic and industrial policy debates.

In order to understand the potential inefficiencies in research investment and to design appropriate industrial policies to address them, it is necessary to identify the spillovers associated with basic research. This paper posits that basic research differs from applied research in two important ways. First, significant advances in scientific knowledge that form the basis for subsequent applied innovations come through basic research. Second, these significant basic advances can potentially be applicable to many different industries.

The role played by Du Pont de Nemours in the financing of William Carothers' research and in the development of its subsequent applications provides a good illustration of these spillovers. According to Nelson (1959), "Carothers' work in linear super-polymers began as an unrestricted foray into the unknown, with no practical objective in mind. But the research was in a new field in chemistry and Du Pont believed that any new chemical breakthrough would likely be of value to the company. In the course of research Carothers obtained some super-polymers that became viscous solids at high temperatures, and the observation was made that filaments could be made from this material if a rod were dipped in the molten polymer and withdrawn. At this discovery the focus of the project shifted to these filaments and Nylon was the result." Nylon is now used in many industries such as textiles, automobiles, and military hardware, three industries in which DuPont had operations.

Indeed the scientific knowledge generated by basic research can potentially be of use in more than a single field (Nelson (1959), Rosenberg (1990), Dasgupta and David (1994)). Ideally, in order to capture the full return from this new scientific knowledge in industries where it could have an application but in which the innovating firm is not present, the innovator would first patent and then license or sell the innovation to other firms in those industries. However, the applications of basic scientific advances are often not immediate and firms are often only able to transform them into patentable applications in their own industries. This is the well-known *approprability problem* of basic research which is initially discussed in Nelson (1959) and again in Rosenberg (1990) and Dasgupta and David (1994). Hence, firms operating in more industries will have a greater probability of being able to directly utilize more facets of a basic innovation. As it was hypothesized in Nelson (1959): "It is for this reason that firms which support research toward the basic-science end of the spectrum are firms that have *fingers in many pies*". Note that the key concept that is being emphasized here is not firm size per se, but the diversity of its operations.

This interesting argument (which we will refer to as Nelson's hypothesis) will be the central focus of our analysis in this paper.

Because of these potentially sizable spillovers both *within* and *across* industries, the consensus has been that there is an underinvestment in basic research, although there is no agreement about the extent of the underinvestment.[2] A satisfactory analysis of the questions listed above requires a structural framework that models the incentives for different types of research investments by private firms. We strongly believe that the combination of a macro model with heterogeneous firms and firm-level data would greatly contribute to our understanding of the study of macro-level policies. Our goal in this project is to take a first step toward developing this theoretical framework, to estimate it using micro-level data, and to discuss the effects of different research policies on productivity and growth. To the best of our knowledge, this would be the first study to model private investment in basic and applied research simultaneously as well as government investment in basic research in an endogenous growth framework and the first that combines a quantitative analysis with micro-level evidence to identify sources of spillovers arising from basic research.

Our analysis proceeds in three steps. We first propose a novel way of empirically identifying spillovers from basic research. Second, motivated by those empirical facts, we propose a general equilibrium, multi-industry framework with private firms and a public research sector. Firms conduct both basic and applied research, whereas the public sector focuses exclusively on basic research.[3] In our model, basic research generates fundamental technological innovations and generates spillovers that affect subsequent applied innovations.[4] In line with the "Ivory Tower" theory of academic research, basic research by private firms in our model will turn into consumer products faster than that undertaken by public research labs (for empirical evidence, see Trajtenberg, Henderson, and Jaffe (1992), Rosenberg and Nelson (1994), Henderson, Jaffe, and Trajtenberg (1998) and more recently Bikard (2013)). Applied research, on the other hand, will be done only by private firms and will generate incremental improvements in the fundamental technologies developed through basic research. To highlight the key economic forces, we will first consider a benchmark economy with tractable functional forms, characterize the dynamic equilibrium, and discuss the resulting dynamics and inefficiencies. Next, we will introduce the general framework. After characterizing the equilibrium of the general model, we estimate the structural parameters. Finally, we use the estimated model to assess the extent of inefficiencies in basic and applied research and to study the implications of several important R&D policies.

[2] Our paper is a contribution to the vast empirical literature on R&D spillovers (see for example, Griliches (1986), Jaffe, Trajtenberg, and Henderson (1993), Audretsch and Feldman (1996), Anselin, Varga, and Acs (1997), Bloom, Schankerman, and Van Reenen (2013)).

[3] Trajtenberg, Henderson, and Jaffe (1992) show that academic research is more basic and less frequently patented.

[4] By fundamental innovation, we mean the major technological improvements that generate larger than average contributions to the aggregate knowledge stock of society. In addition, these will have long-lasting spillover effects on the size of subsequent innovations within the same field.

The reduced-form analysis of our paper contributes to the empirical innovation literature by introducing two new ways of identifying spillovers (see Bloom, Schankerman, and Van Reenen (2013) for a recent literature review). First, we use the variation in the level of basic research spending between firms operating in different numbers of industries to infer the magnitude of cross-industry spillovers. Second, we use heterogeneous citation patterns across public and private patents in order to identify within-industry spillovers.

The theoretical framework directly contributes to the literature on endogenous growth. Although the different characteristics of basic and applied research and public and private research have been widely recognized to be of first-order importance by policymakers, these issues have received insufficient attention from the economic growth literature. In particular, models of endogenous technological change (see Aghion, Akcigit, and Howitt (2013) for a recent survey) mainly considered a uniform type of (applied) research and ignored basic research investment in the economy. A few exceptions, such as Aghion and Howitt (2009), Morales (2004), and Mansfield (1995), have considered theoretical models where basic research is done publicly but ignored the private investment in it. In addition, we enrich the analysis of the distinct features of basic research by identifying within- and cross-industry spillovers.

Quantitatively, our contribution is to consider the distinct spillovers associated with basic research, model private firms' incentives to invest in basic research and hence quantify the size of the inefficiencies in the decentralized economy. Our analysis suggests that the size of the inefficiency is around 4.7 percentage points in consumption equivalent terms, which raises an important question: to what extent can public policies address this inefficiency?

Our final contribution is to the growing literature on the quantitative assessment of the effect of innovation and industrial policies on productivity and welfare (see, for instance, Acemoglu, Akcigit, Bloom, and Kerr (2013), Garicano, LeLarge, and Van Reenen (2013), Impullitti (2010), among others). We first evaluate the impact of innovation policies considered by policymakers in many OECD countries. The first policy we analyze is a uniform research subsidy to private firms. Various papers have empirically shown that R&D subsidy policy has been ineffective (Romer (2001), Goolsbee (1998), and Wilson (2009)). This result is theoretically puzzling as standard endogenous growth models typically predict that growth rises when private R&D is subsidizied at the expense of lower initial consumption (see Aghion and Howitt (1998) pp. 486 and Acemoglu (2008) pp. 478 for more detailed discussions). These models feature only a single type of research. Once the distinction between basic and applied research is introduced, the results can differ greatly and shed light on the aformentioned puzzle. We show that in an economy with both types of research, the major underinvestment is in basic research due to its sizable spillovers. In this environment, subsidizing overall private research is less effective since this policy oversubsidizes applied research, which is already overinvested in due to competition. Therefore, the welfare improvement from such a subsidy is limited unless the policymaker is able to discriminate between types of research projects at the firm level, which is considered to be quite impractical. Nevertheless, we consider a hypothetical type-dependent research subsidy and find that the optimal policy is to subsidize basic research by 50% and applied research by 14%.

We then analyze the optimal allocation of funding to public research labs. Due to the Ivory Tower nature of public basic research, allocating more money to the academic sector without giving property rights to the researcher is not necessarily a good idea. Therefore, we mimic a policy exercise similar to the Bayh-Dole Act enacted in the US in 1980. We consider alternative scenarios in which the public researcher has no property rights, then 50% and 100% property rights. We find a complementarity between the level of property rights and the optimal allocation of resources to academic research. The optimal combination turns out to be granting full property rights to the academic researcher and allocating 3.7% of GDP to public research. This reduces the welfare gap to from 4.7 to 1.7% in consumption equivalent terms.

The rest of the paper is organized as follows. Section 2 introduces some new empirical facts on basic research spillovers to motivate our modeling approach. The discussion of our theoretical framework consists of two parts: In Section 3.1 we provide a benchmark version of the main model, characterize its dynamic equilibrium in an intuitive manner and discuss the main mechanisms present in the model. In Section 4 we describe a generalization of the benchmark model that we bring to the data and estimate. Section 5 describes our quantitative analysis, including the welfare properties of the estimated model. Section 6 provides a detailed discussion of the welfare effects of various policies on the decentralized economy. Section 7 concludes. The Appendix contains omitted proofs and derivations (A), the data description (B), further theoretical details (OA-1), further robustness checks on the stylized facts (OA-2), further details on within-industry spillovers (OA-3), and target moments and identification (OA-4).

2 Empirical Facts

Despite the extensive literature in the fields of industrial organization and endogenous growth focusing on aggregate research investment, the literature on firms' basic research decisions has been very thin and empirical studies are even rarer. Notable exceptions include Mansfield (1980, 1981), Griliches (1986) and Link (1981), who empirically document the positive contribution of basic research to firms' productivity. Existing studies on basic research have been mainly theoretical and have devoted their attention to academic/public research as the source of basic research, ignoring the *private* sources of it (Segerstrom (1998); Aghion and Howitt (1996), Morales (2004)). We believe that part of the reason for this outcome, as argued in one congressional report (JEC (2010)), has been the lack of firm-level data on distinct types of private research expenditures. Our empirical evidence also contributes to the literature on innovating firms (e.g., Klette and Kortum (2004), Lentz and Mortensen (2008), Akcigit and Kerr (2010), and Acemoglu, Akcigit, Bloom, and Kerr (2013)), characterizing the innovation rates and size dynamics of firms conducting R&D.

In this paper we use unique data on the French economy combining information not only on product market and R&D investment characteristics of individual firms but also on plant and ownership information for the period 2000-2006. The R&D information comes from an annual survey conducted by the French Ministry of Research that covers a large, representative cross-section of innovating French firms. In this survey firms are asked to report their expenditures for basic and applied research.[5] Details regarding data sources are provided in Appendix B.

The next section presents the main empirical facts emerging from these data.

2.1 Basic Versus Applied Research

First we document that private firms' investment in basic research forms a non-negligible fraction of both total private research spending and total basic research spending. Table 1 reports official statistics from the French Ministry of Research on public investment in basic research and private investment in applied and basic research for the period 2000-2006.

Table 1: Expenditures On Basic Research

	Private		Public	Private	Basic
Year	Basic	Applied	Basic	$\left(\frac{\text{Basic}}{\text{Applied}}\right)$	$\left(\frac{\text{Private}}{\text{Public}}\right)$
2000	802	7005	6425	.11	.13
2001	795	7748	6786	.1	.12
2002	959	8899	7037	.11	.14
2003	1092	8928	7133	.12	.15
2004	1175	9482	7338	.12	.16
2005	1227	9469	7331	.13	.17
2006	1213	10278	7755	.12	.16

Notes: Expenditures on basic and applied research in millions of euros for the period 2000-2006. Source: French Ministry of Research.

Private spending on basic research amounted to an average of 1 billion euros per year as opposed to 8.3 billion on applied research for the period 2000-2006. During the same period, public expenditures on basic research represented an average of 7 billion euros per year in France.[6] This implies that more than 11% of private research is spent on basic research. More important, almost 15% of total basic research in the economy is undertaken by private entities.[7]

The picture that emerges therefore hints at a significant involvement of the private sector in undertaking basic scientific research. Thus, ignoring the private incentives behind basic research might prevent economists and policymakers from designing more effective policies for productivity growth.

[5] The definition used by the French authorities for basic and applied research is based on the Frascati manual. It is therefore similar to the NSF definition presented in the introduction.

[6] Public research has three major components: public research labs, universities, and the French National Science Foundation (CNRS). Their relative shares within public research expenditures are around 20%, 40% and 40%, respectively. To simplify the terminology we will refer to them as public research labs.

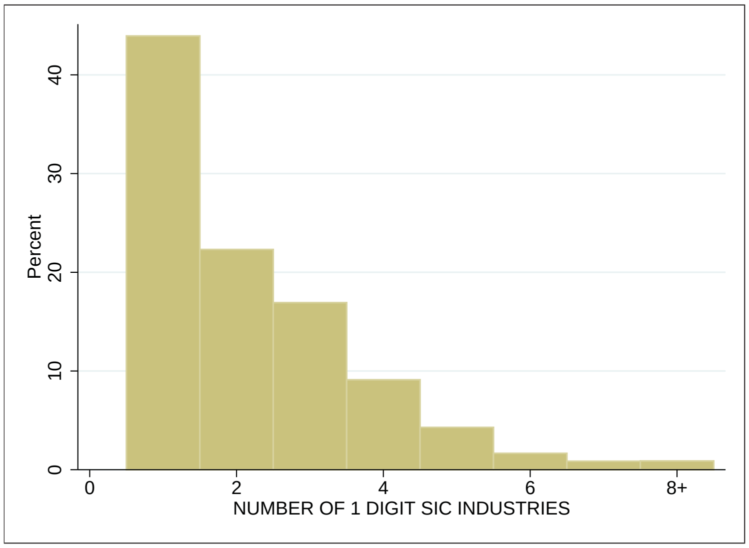
[7] Similarly, Howitt (2000), using an NSF survey, finds that around 22% of all basic research in the US during the period 1993-1997 was performed by private enterprises.

2.2 Multi-Industry Distributions

Another stylized fact emerging from the data is the extent of the multi-industry presence of firms. Figure 3 uses our micro-level data on French companies between 2000 and 2006 in order to plot their empirical distribution into multiple industries.

To measure multi-industry presence, we count the number of distinct SIC codes in which a firm is present. Our data allow us to identify a firm's links to different industries not only through product lines within the same firm but also through its majority ownership links. To avoid misclassification of related industries, we consider as our benchmark case the number of distinct 1-digit SIC codes (10 industries).

Figure 3: Distribution of Firms by Number of Industries



On average firms are present in 2 distinct industries as defined by 1-digit SIC codes. Although nearly 44% of the firms are operating in only one industry, the remaining firms occupy a large spectrum of industries. Insights from Figure 3 are very similar when using more disaggregate SIC classifications (up to the 4-digit SIC level) or when changing the definition of an industry link.[8]

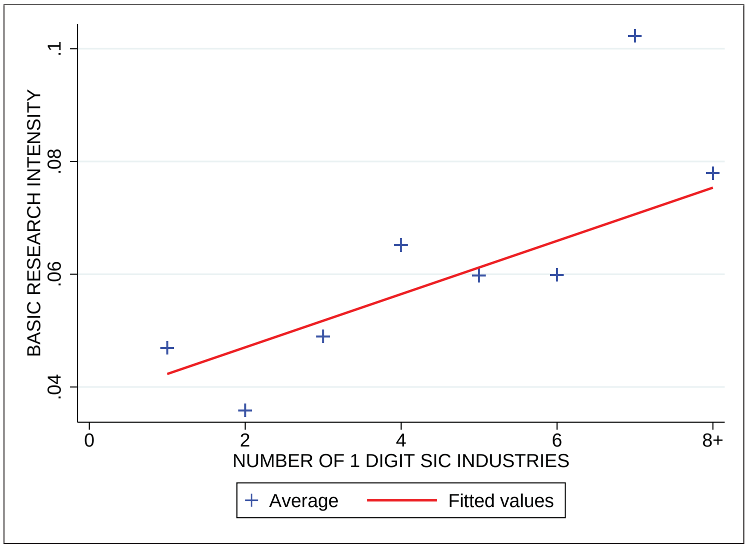
2.3 Basic Research and Cross-Industry Spillovers

Next we link our stylized fact on multi-industry presence to private incentives for basic research. More specifically, we test Nelson's hypothesis that the main investors in basic research would be those firms that have *fingers in many pies*. According to this argument, as the range of a firm's products and industries gets more diversified, its incentive for investing in basic research relative to applied research should increase due to better appropriability of potential knowledge spillovers.

[8]Figures available upon request.

Figure 4 plots average basic research intensity against the total number of distinct 1-digit SIC codes in which the firm is present together with a simple linear fit of the data. Basic research intensity is defined as the ratio of total firm investment in basic research to total firm investment in applied research.

Figure 4: Basic Research Intensity and Number of Industries



Note: Linear plot on pooled data for the period 2000-2006. *Basic Research Intensity* is defined as the ratio of total firm investment in basic research divided by total firm investment in applied research. *Number of 1-Digit SIC Industries* is the number of distinct SIC codes in which a firm is present.

Figure 4 shows that a simple linear fit of the data suggests a positive and statistically significant relationship between the two variables. Table 2 provides further evidence about the relationship between multi-industry presence and basic research intensity. To take into account the corner solution at 0, we estimate the relation between a firm's basic research intensity and its multi-industry presence using a Tobit model.

In all specifications basic research intensity is increasing in the number of industries. According to the benchmark estimation, presence in an additional industry increases a firm's basic research intensity by 3 percentage points on average. In terms of magnitude, this corresponds to a 50% increase in the average research intensity of a single industry firm. Note also that the magnitude of the estimated coefficient on multi-industry presence is stronger for higher levels of SIC aggregation (for less related activities), since it decreases from 3 percentage points at the 1-digit SIC level to 2.1 percentage points at the 4-digit SIC level.

Table OA-1 in Section OA-2 of the online appendix provides a rich set of robustness checks in terms of control variables, alternative measures of multi-industry presence, and estimation methods. Most important, it exploits historical ownership structures and changes in government policies as instrumental variables. The IV estimates are larger in magnitude and seem to suggest that the positive correlation is not driven by omitted variables.

Table 2: Basic Research Intensity and Multi-Market Activity

	1-Digit SIC	2-Digit SIC	3-Digit SIC	4-Digit SIC
Log # of Industries	0.032***	0.027***	0.024***	0.021***
	(0.01)	(0.00)	(0.00)	(0.00)
Log Employment	0.003**	0.002	0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)
Year & Organization Fixed Effects	YES	YES	YES	YES
N	13708	13708	13708	13708

Notes: Pooled data for the period 2000-2006. Estimates are obtained using Tobit models and relate to the marginal effect of the regressors at the sample mean. *Basic Research Intensity* is defined as the ratio of total firm investment in basic research divided by total firm investment in applied research. *Log # of Industries* is the number of distinct SIC codes in which a firm is present. Robust standard errors clustered at the firm level in parentheses. See appendix for the definition of variables. One star denotes significance at the 10% level, two stars denote significance at the 5% level, and three stars denote significance at the 1% level.

2.4 Basic Research and Within-Industry Spillovers

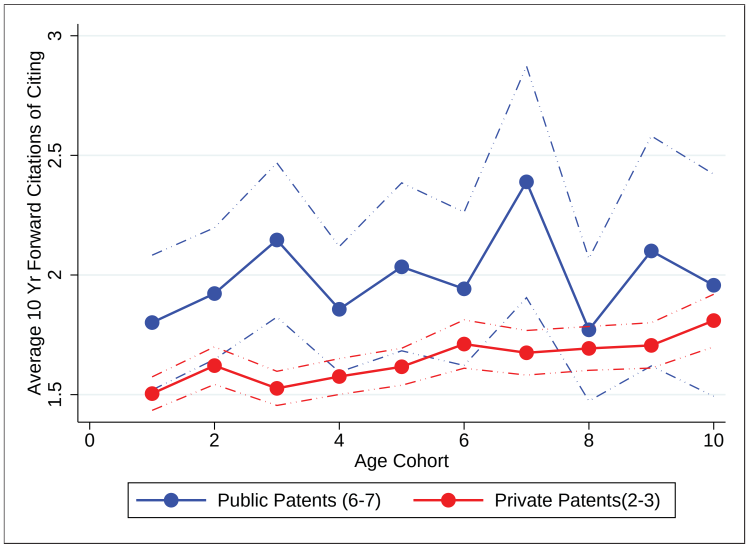
Basic research contributes to economic growth through its impact on subsequent innovations within the same industry as well. Applied research builds on the latest technological knowledge in the product line but the returns from building on the original breakthrough innovation diminishes as more and more firms exploit it.

To empirically capture these spillovers, we turn to patent data. The idea is to pin down the age at which a patent derived from basic research cannot be distinguished from a patent derived from applied research in terms of its importance for follow-up innovations. Two empirical issues need to be addressed: (*i*) distinguishing patents derived from basic and applied research, and (*ii*) capturing the idea of successively less original contributions. We address the first point by distinguishing between patents applied for by corporations from patents applied for by public institutions. We address the second point by computing a citation-based measure of the marginal contribution of citing patents over time.

To do so, we use the NBER patent data set covering the period 1974-2006. The analysis of our final data set will focus exclusively on French patentors but the construction of the different variables uses information from the entire data set. For each patent we first identify citing patents across time. The age of a patent is given by the difference between the grant year of the patent and the current year. For each of the citing patents we compute the cumulative 10-years-forward citations these citing patents receive. For each originally cited patent we are then able to compute across age the mean of the citing patents' cumulative 10-years-forward citations. Our measure, *Average Citations of Citing Patents*, captures the marginal importance of each successive citing patent. Appendix OA-3 provides further explanations of the construction of this variable.

Figure 5 presents graphical evidence for French patents between 1975 and 1985. It plots *Average Citations of Citing Patents* for French public patents (blue line) and French private patents (red line).

FIGURE 5: CITATION PATTERNS FOR FRENCH PUBLIC AND PRIVATE PATENTS



*Note:*Panel plots *Average Citations of Citing Patents* for French public patents (blue line) and French private patents (red line) across patent age. *Average Citations of Citing Patents* is computed as the 10-years-forward citations of the citing patents and is measured for patents granted in the period 1975-1985.

The figure measures *Average Citations of Citing Patents* computing the 10-years-forward citations of the citing patents. Patents citing private patents receive on average 1.6 citations within the first 10 years. The relative importance of patents citing private patents remains stable and slightly increasing through the age of the private patent. Patents citing public patents receive on average two citations within their first 10 years. The importance of citing patents is stable until the original patent is 8 years old. At this moment we observe a significant drop in citations of citing patents from 2.4 to 1.7. This is when the difference between private and public, in terms of citings' citations, becomes non-significant. Although public citations of citing patents slightly increase again after this drop, the difference remains smaller and

statistically non-significant, as indicated in Table 3. The results are similar when using the Wilcoxon-Mann-Whitney test. Appendix OA-3 provides further robustness checks related to the computation of the citations variable and related to the public and private patent classification.

Our results are consistent with previous stylized facts related to citations of private and academic patents. Henderson, Jaffe, and Trajtenberg (1998) as well as Trajtenberg, Henderson, and Jaffe (1997) show that corporate patents relative to academic patents tend to be relatively less cited and less general in the subsequent technological fields that cite them. Evidence on European patent data is more scarce, but Bacchiocchi and Montobbio (2009) show that university patents and public research organizations are more highly cited during the first five years but then become similar in

TABLE 3: CITATION DIFFERENCES FOR FRENCH PUBLIC AND PRIVATE PATENTS

Age	1	2	3	4	5	6	7	8	9	10
Difference	.3**	.3**	.62***	.28**	.41**	.23	.71***	.08	.39	.14
	(0.15)	(0.15)	(0.17)	(0.14)	(0.18)	(0.17)	(0.25)	(0.16)	(0.25)	(0.24)

Notes: Differences in citation patterns of 15383 patents granted by the USPTO to French private (92%) and public (8%) depositors. The difference is computed in terms of *Average Citations of Citing Patents* across patent age. *Average Citations of Citing Patents* is computed as the 10-years-forward citations of the citing patents and is measured for patents granted in the period 1975-1985. Two sample t-test with unequal variances were used. One star denotes significance at the 10% level, two stars denote significance at the 5% level, and three stars denote significance at the 1% level.

terms of citations.

3 Theory: Growth with Basic and Applied Research

We first discuss important choices in modeling before going into specific details. Our theoretical framework will depart from standard endogenous growth models in a number of ways. In line with stylized Fact 1, we will allow private firms to invest in both basic and applied research. Second, to capture stylized Fact 2, firms will be able to operate in multiple industries. Third, our analysis relies on the appropriation of spillovers from basic research by multi-industry firms; hence, there will be cross-industry spillovers from basic research. Fourth, to capture stylized Fact 4, there will also be within-industry spillovers from basic research. Finally, we will also introduce a public research sector, which can be thought of as universities or publicly funded research labs. The key distinction between private basic research and public basic research will be that an outcome of the former will turn immediately into a consumer product of the innovating firm, while the latter will contribute to the general pool of basic knowledge and will not turn into a consumer product until a firm uses that knowledge. This will induce a delay in the effect of public basic research, as argued in the introduction. The social trade-off will be that while private firms are better at turning abstract basic research into consumer products, they do not internalize all the spillovers associated with it. Hence, there will be room for meaningful policy interventions, which we will provide after our quantitative analysis.

For ease of exposition and intuition, in this section, we will first outline a simplified baseline framework with myopic (one-period-ahead maximizing) firms that highlights the key elements of the main model. After deriving the theoretical results and discussing the main economic forces, in Section 4 we will describe the generalizations we make to the benchmark model.

3.1 Baseline Model

We consider a representative household economy in continuous time. The household consists of a measure M of workers. Each worker has one unit of labor that is supplied inelastically in the labor market. There is a unique final good $Z(t)$. The economy is a closed economy, there is no physical capital investment and all expenses are in terms of the labor units; therefore, $Z(t)$ will also be equal to household consumption at time t.

3.1.1 Production

Production is divided into three major sectors: *downstream*, *midstream,* and *upstream* sectors. The upstream sector produces intermediate goods (y_{ij}) that are used to produce industry aggregates (Y_i) in the midstream sector. Finally, the downstream sector combines these industry aggregates into the final good (Z). We now describe them in detail.

Downstream Sector Final good $Z(t)$ is produced in the downstream sector by infinitely many competitive firms that combine inputs from M different industries according to the following constant elasticity of substitution (CES) production function

$$Z(t) = \left[\frac{1}{M}\sum_{i=1}^{M} Y_i(t)^{\frac{\gamma-1}{\gamma}}\right]^{\frac{\gamma}{\gamma-1}}. \tag{1}$$

In this production function, $Y_i(t)$ is the aggregate output from industry $i \in \{1, ..., M\}$. The economy consists of $M \in \mathbb{Z}_+$ industries. In the context of firm-level data, each industry i can be thought of as a different 1-digit Standard Industrial Classification (SIC) code and $Z(t)$ is simply the aggregate GDP of the economy.[9] We normalize the price of the final good to 1 at every instant t without any loss of generality. For notational simplicity, time subscripts will henceforth be suppressed.

Midstream Sector Each industry aggregate Y_i is produced competitively, combining inputs from a continuum of product lines. Let y_{ij} denote the production of upstream good j in industry i by the firm that has the best technology in that product line. Industry aggregate i is produced according to the following CES production function

$$Y_i = \left[\int_0^1 y_{ij}^{\frac{\varepsilon-1}{\varepsilon}} dj\right]^{\frac{\varepsilon}{\varepsilon-1}}. \tag{2}$$

Upstream Sector In product line j, the firm that has the latest (and also the best) technology has the monopoly power and produces according to the following linear production technology that takes only labor as an input

$$y_{ij} = q_{ij} l_{ij} \tag{3}$$

where $q_{ij} > 0$ is the labor productivity associated with product line j and l_{ij} is the number of production workers employed. Let us denote the wage rate in the economy by w in terms of the

[9]Note that we introduce this multi-industry structure in order to model cross-industry spillovers. To avoid any additional theoretical complications, we will focus on symmetric equilibria in which industry aggregates assume a common value.

final good. The specification in (3) implies that each product y_{ij} has a constant marginal cost of production $w/q_{ij} > 0$. We denote the productivity index of industry i by

$$\bar{q}_i \equiv \left(\int q_{ij}^{\varepsilon-1} dj \right)^{\frac{1}{\varepsilon-1}}. \tag{4}$$

Definition of a Firm In this model, as in Klette and Kortum (2004), a firm is defined as a collection of product lines in which it is the lead producer. These product lines can come from multiple industries. In what follows, $m_f \in \{1, .., M\}$ will denote the number of industries in which the firm actively operates, $n_{if} \in \mathbb{Z}_+$ will denote the number of product lines firm f owns in a given industry i, and finally n_f will stand for the total number of product lines of the firm such that $n_f \equiv \sum_{i \in m} n_{if}$. For notational tractability, henceforth we will drop the firm index f, when it creates no confusion.

A firm's payoff in a given product line j in industry i depends on its productivity level q_{ij}. Therefore, the payoff-relevant state of a firm is denoted by

$$\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m)$$

where $\mathbf{q}_i = \{q_{i,1}, q_{i,2}, \ldots, q_{i,n_i}\}$ is a multi-set keeping track of all the productivity levels of the firm in industry i where it has the best technology.[10] Working with such a large and complex state space proves burdensome in practice. Later on, we will impose sufficient assumptions that allow us to use a much simpler equivalent representation for a firm's product portfolio.

Example 1 *An example is helpful to summarize the description so far. Figure 6 illustrates an example of an economy that consists of $M = 3$ industries. It also shows an example of a firm (f) that operates in $m = 2$ industries ($i = 1$ and $i = 3$) and has $n_1 = 3$ product lines in industry $i = 1$ and $n_3 = 2$ product lines in $i = 3$. This firm does not currently operate in industry $i = 2$.*

A firm's portfolio of products will expand through successful innovation. Likewise, it will lose product lines when other firms or potential entrants successfully innovate on one of its product lines (thus stealing it). These innovations will be the source of economic growth in this economy. The next subsection will describe the details of the innovation technology.

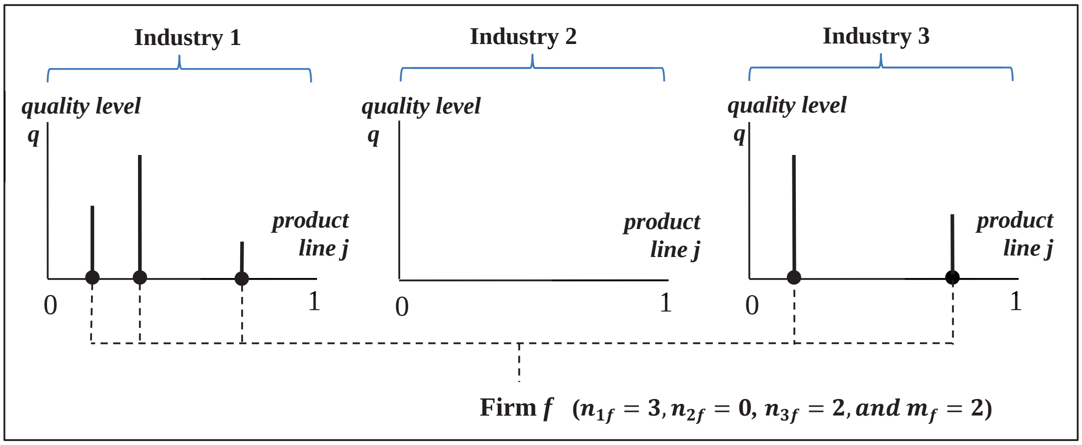
3.1.2 Innovation and Technological Progress

In this economy, there are two types of innovations (basic and applied) and two different groups of agents (private and public sectors) generating productivity growth.

Firms innovate by investing in two types of research: *basic* and *applied*. In accordance with Nelson (1959)'s description, significant advances in technological knowledge come through basic innovation in our model. Applied innovation builds on the existing basic innovations. Innovations run

[10] A multi-set is a generalization of a set that can contain more than one instances of the same member. For instance, given $j \neq j'$, a multiset $\mathbf{q}_{if}$ can contain $q_{if}(j)$ and $q_{if}(j')$ regardless of whether $q_{if}(j) = q_{if}(j')$.

Figure 6: Example of a firm



into diminishing returns. If the latest basic innovation in a product line becomes outdated, applied innovations in that product line become less productive until a new basic innovation introduces additional fundamental knowledge that can make the applied innovation more productive again. Therefore, there will be complementarity between the two types of innovation at the aggregate level.

These innovations come from two sources: First, the private sector invests in both basic and applied innovation with the goal of increasing their market share. Second, the government uses tax revenues to fund public research labs to produce basic innovations. In what follows, we are going to describe firms' research technology and the distinction between basic and applied research. Then we will describe the public research technology.

Research by Private Firms Firms choose their flow rate of innovation and pay a labor cost that is increasing and convex in this rate. Basic and applied research levels are chosen separately, and there is no complementarity between them in terms of research costs. In terms of the innovation production function, we will follow the literature (see Klette and Kortum (2004), Lentz and Mortensen (2008), Acemoglu, Akcigit, Hanley, and Kerr (2012), and Acemoglu, Akcigit, Bloom, and Kerr (2013)). Firms undertake innovation by combining their existing, non-tradable intangible capital with researchers (hired at wage rate w as with the production workers) in a Cobb-Douglas production function. In our model, the intangible capital stock in a particular industry i is proxied by the number of product lines n_i that a firm owns in that industry. The production function for applied and basic research then takes the following form

$$A_i = n_i^{1-\frac{1}{\nu_a}} H_{ai}^{\frac{1}{\nu_a}} \Omega_a \quad \text{and} \quad B_i = n_i^{1-\frac{1}{\nu_b}} H_{bi}^{\frac{1}{\nu_b}} \Omega_b$$

where $\Omega_a, \Omega_b > 0$ are scale parameters, $\nu_a, \nu_b > 1$ are the inverse of the innovation production function elasticities with respect to researchers and H_{ai} and H_{bi} denote the number of researchers that firm f needs to hire in order to generate the Poisson flow rates for applied (A_i) and basic research (B_i) in industry i.

The above specifications, which are standard in this class of models, capture the idea that a firm's knowledge capital facilitates innovation.[11] Let us define $a_i \equiv A_i/n_i$ and $b_i \equiv B_i/n_i$ as the applied and basic *innovation intensities*. Similarly, let $h_a(a_i) \equiv H_{ai}/n_i$ and $h_b(b_i) \equiv H_{bi}/n_i$ be defined as the number of researchers per product line hired for applied and basic research. As a result, we can summarize the cost of doing applied and basic research as

$$C_a\left(a_i \mid n_i\right) = w n_i a_i^{\nu_a} \xi_a \quad \text{and} \quad C_b\left(b_i \mid n_i\right) = w n_i b_i^{\nu_b} \xi_b \tag{5}$$

where w is the wage rate, $\xi_a \equiv \Omega_a^{-\nu_a}$, and $\xi_b \equiv \Omega_b^{-\nu_b}$. Notice that total cost is directly proportional to the number of product lines.

Similar to Klette and Kortum (2004), Lentz and Mortensen (2008), Acemoglu, Akcigit, Hanley, and Kerr (2012), and Acemoglu, Akcigit, Bloom, and Kerr (2013), both applied and basic research are *directed* toward particular industries but *undirected* within those industries. In other words, once a firm chooses A_i and B_i, the realization of innovations will take place on a random product within industry i.

Innovation through *basic research* introduces a new generation of fundamental technical knowledge. The utilization of this fundamental knowledge for production requires what we call industry-specific *working knowledge*. This translates into our model one of the main insights on basic research presented by Nelson (1959). Although the knowledge generated by basic research is often applicable to many industries, the ability to turn it into patents and capture its full economic value critically depends on the spectrum of activities and technologies operated by the firm. In the model, each firm has therefore some working knowledge in the industries where it has undertaken production (m). For now, we take the joint distribution $\Gamma_{m,n}$ over m and n as given but we will endogenize it in the generalized model in Section 4.

Let $q_{ij}(t)$ be the highest productivity technology for producing j in industry i. When a firm that has working knowledge in i produces a basic innovation that has a direct application in industry i and product line j, the same firm uses this basic knowledge for production and patents this new high-value technology. As a result, the firm improves $q_{ij}(t)$ by $\eta \bar{q}_i(t)$

$$q_{ij}\left(t + \Delta t\right) = q_{ij}\left(t\right) + \eta \bar{q}_i(t) \tag{6}$$

where $\eta > 0$ is the step size, and $\bar{q}_i$ is the productivity index defined in equation (4). When the firm produces this new innovation, it adds this product line with the productivity improvement into its portfolio $\mathbf{q}(t + \Delta t) = \mathbf{q}(t) \cup \{q_{ij}(t + \Delta t)\}$, which generates per-period profit of $\pi\left(q_{ij}(t + \Delta t)\right)$. Going back to Example 1, firm f would increase its total number of product lines from 5 to 6 with this basic innovation.

Moreover, basic research features two potential spillovers:

[11] It also simplifies the analysis by making the problem proportional to the number of product lines.

- *within-industry spillover (Fact 2.4)*: Each new basic innovation changes the evolution of the product line by introducing a radically new technology. The introduction of this new basic technology causes subsequent applied innovations to be larger until the latest basic technology becomes outdated through some random process. We refer to product lines just hit by basic innovation as *hot* product lines, as opposed to *cold* product lines, whose latest basic innovation has become outdated.

- *cross-industry spillover (Fact 2.3)*: Each new basic innovation has the potential for spillovers into other industries. With some probability a basic innovation will generate an additional basic innovation in some other industry. If the firm has working knowledge in this other industry, it can use the innovation for production. Otherwise, the new technology contributes to the pool of existing basic knowledge and will eventually contribute to a new consumer product made by some other producer.

These two types of spillovers lie at the heart of our analysis; therefore, we will now discuss each in more detail.

Within-Industry Spillover from Basic Research *Applied research* makes use of the *within-industry* spillover from basic research and builds on the existing latest basic technological knowledge in a product line. The productivity of each applied innovation is a function of how depreciated the latest basic technology is. If the latest basic knowledge in j is undepreciated (i.e., still hot), a successful applied innovation will benefit from it and improve the latest productivity $q_{ij}(t)$ of that product line by $\eta\bar{q}_i(t)$, as in expression (6): $q_{ij}(t+\Delta t) = q_{ij}(t)+\eta\bar{q}_i(t)$. If the latest basic technology of the product line is depreciated (i.e., cold), a successful applied innovation will improve the latest productivity only by a magnitude proportional to $\lambda < \eta$ so that

$$q_{ij}(t + \Delta t) = q_{ij}(t) + \lambda\bar{q}_i(t). \tag{7}$$

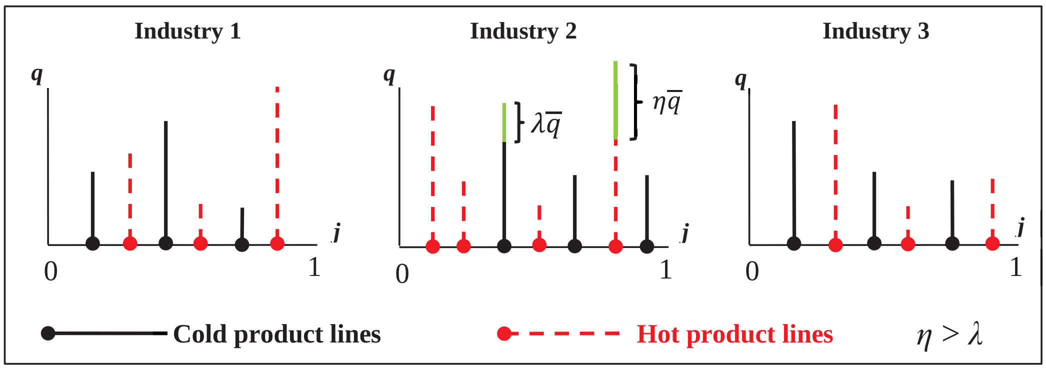
We assume that a new basic technology depreciates (innovations run into diminishing returns) at a Poisson rate $\zeta > 0$. On the other hand, a new basic innovation resets the product line until the next time it cools down again. Let us denote the arrival rate of basic innovations to product lines by τ_b. Then during a small time interval Δt, each product line will be subject to the transition rates denoted in Table 4:

Table 4: Transition matrix for within-industry spillovers

	hot	**cold**
hot	$1 - \zeta\Delta t$	$\zeta\Delta t$
cold	$\tau_b\Delta t$	$1 - \tau_b\Delta t$

Figure 7 illustrates the implications of within-industry spillovers. In every industry, at any point in time, some product lines will be hot (red dotted lines) and some product lines will be cold, in cases where the latest technology is outdated (black solid lines). We will denote the share of hot product lines by $\alpha_i \in [0,1]$. In a balanced-growth-path equilibrium, the share of the hot product lines will be determined through the transition rates in Table 4 and will remain invariant. An applied innovation is more productive if the latest basic knowledge in that product line is still "hot" and improves the productivity by $\eta\bar{q}_i$; otherwise, the contribution is only $\lambda\bar{q}_j$ where $\eta > \lambda$. This highlights the complementarity between basic and applied research.

Figure 7: Within-industry spillover



Cross-Industry Spillover from Basic Research Basic research features an additional element of uncertainty arising from random spillovers into other industries. When a firm successfully innovates through basic research, the resulting new fundamental knowledge will be applied first by that firm to increment the productivity of a random product in the target industry.

The characteristic feature of basic research we wish to capture is that it often has applications in many industries other than the one for which it was originally intended (Nelson (1959)). Therefore, we will assume that when a basic innovation occurs, it applies with probability one to the target industry, and with probability $p \in (0,1)$, it generates an additional basic innovation in another industry determined by nature at random. Thus, p is our measure of the intensity of cross-industry spillovers. Let $\mathbf{1}_{i,i'}$ be an indicator function that takes a value of one if a basic innovation in industry i has an application in industry i' and zero otherwise. Then the unconditional probabilities satisfy

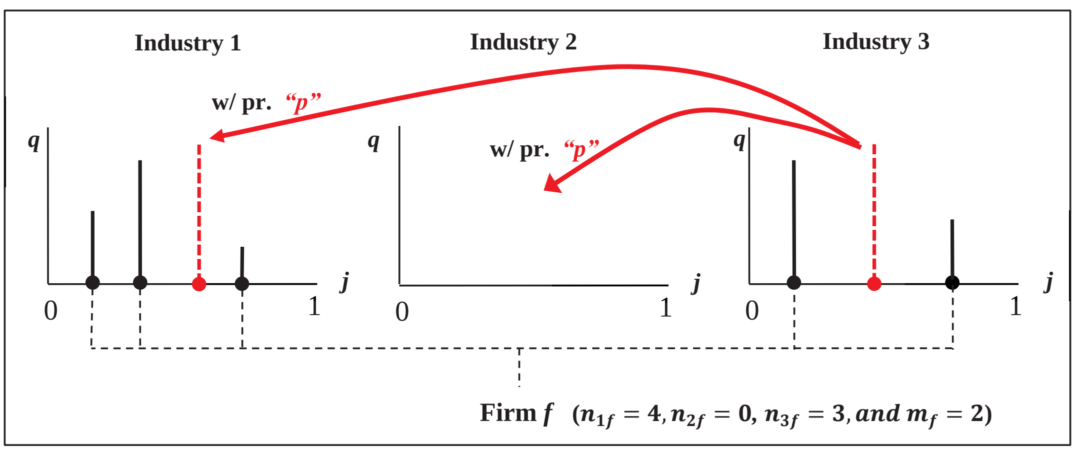
$$\Pr\left[\mathbf{1}_{i,i'} = 1\right] = \begin{cases} \frac{p}{M-1} & \text{if } i' \neq i \\ 1 & \text{if } i' = i \end{cases}. \tag{8}$$

The spillover innovation in industry i' will be of step size η as well but will not generate additional cross-industry spillovers. This new innovation will be used by the same firm f if it has working knowledge in i'. Otherwise the production potential of this innovation will be used by the next inventor in that product line.

This structure captures Nelson's hypothesis. When a firm generates some basic knowledge, it can turn this into an immediate application only in the sectors in which it has working knowledge. Nelson (1959) observes that in order to capture the full return from new basic scientific knowledge in industries where a firm is not present but the knowledge could have an application, the innovating firm must first patent and then license or sell the innovation to other firms in those industries. However, the applications of significant scientific advances are often *not immediate* and firms can turn them into patentable applications mostly in their own industries due to their expertise in the field.

Example 2 *Cross-industry spillovers are depicted in Figure 8. Firm f from Example 1 now produces a basic innovation in industry 3. This adds a new product line to the firm's portfolio and hence the number of product lines of the firm goes from 2 to 3 in $i = 3$. In addition, this basic knowledge has a potential application in industries $i = 1$ and $i = 2$ with probability p. The spillover in industry $i = 1$ is used by the firm since it has working knowledge there. However, the application in $i = 2$ is not immediate to the firm due to lack of working knowledge and therefore it is not used by the current firm but contributes to the pool of basic knowledge in $i = 2$, which can be used by another firm in the future.*

Figure 8: Cross-industry spillover



Recall that m denotes the number of industries in which a firm has working knowledge. Then the probability of a used spillover for the firm is

$$\rho_m \equiv \frac{p\,(m-1)}{M-1} \in [0,1).$$

This highlights the well-known *appropriability problem* of basic research. There is a significant chance that the new basic knowledge will be relevant to multiple industries, but it is not always clear that a firm will be in a position to exploit all of these avenues of production and patenting.

However, firms operating in more industries will have a greater probability of being able to directly use all facets of a basic innovation. As Nelson puts it, firms that have *fingers in many pies* have a higher probability of using the results of basic research. A broad technological base increases the probability of benefiting from successful basic research.

Public Basic Research In our model, the academic sector will be the other source of basic knowledge creation. One of the main tasks of public research labs in an economy is to produce the necessary basic scientific knowledge that will be part of the engine for subsequent applied innovations and growth. We assume that the public research sector consists of a measure U of research labs *per industry*. Each lab receives the same transfer $\bar{R}$ from the government to finance its research which results in an overall funding level of $R = \bar{R} \times U \times M$.

We assume that each public research lab generates a flow rate of u by hiring h_u researchers with the same basic research technology as a one-product firm in (5), so that $u = \Omega_b h_u^{\frac{1}{\nu_b}}$.[12] This specification implies that the government can affect the basic knowledge pool in the economy through the amount of funds R allocated to the academic sector. The flow rate of basic innovation from the academic sector will satisfy

$$u = \left(\bar{R}/w\right)^{\frac{1}{\nu_B}} \Omega_b \tag{9}$$

where u is the academic basic innovation flow *per lab*. In this economy, R is a policy lever controlled by the policymaker. As with private firms, each basic innovation generated by the academic sector applies to industry i and a random product line j and makes that product line hot. However, this innovation by public labs will turn into output only upon a subsequent private applied innovation. In addition to i, the same basic knowledge will contribute to the basic knowledge pool in another industry $i' \neq i$ and line j' with probability $p \in (0, 1)$. Note that the equilibrium fraction of hot product lines α will be determined by the aggregate rates of public (u) and private (b_m) basic research as well as the cool-down rate (ζ).

Remark It is important to note that we follow the empirical Ivory Tower nature of basic research and assume that innovation done by public labs is turned into consumer products only upon subsequent innovation by private firms. The lag between the creation of publicly funded innovations and actual goods production is empirically shown in a large literature: Trajtenberg, Henderson, and Jaffe (1992), Rosenberg and Nelson (1994), Henderson, Jaffe, and Trajtenberg (1998) and more recently Bikard (2013). This important issue is generally overlooked in the theoretical growth literature. Inclusion of this feature generates some new and interesting dynamics, such as the importance of involvement of the private sector in basic research.

[12] In reality, public research labs may have a different research technology than private labs. However, obtaining data on both the inputs and outputs of individual public labs is difficult. The separate estimation of public and private innovation production functions is left for future research.

Entry and Exit The research technology for a single outside entrant is assumed to be the same as that of applied innovation for a firm with a single product line. Thus if an outside entrant hires h_e researchers, it produces a flow probability of entry of $a_e = h_e^{\frac{1}{\nu_a}} \Omega_a$.

There is a mass E of outside entrants per industry. Varying this parameter will control the relative importance of outside entry in the economy. This will imply that creative destruction arising from new entrants will be equal to $E \times a_e$.

In our model, there will be both endogenous and exogenous channels for firm exit. First, a firm that loses all of its product lines to other competitors will have a value of zero and thus will exit. Second, each firm has an exogenous death rate $\kappa > 0$. When this occurs, the firm sells all of its product lines to random firms at a "fire sale" price $\mathcal{P}$.[13] On the flip side, firms will receive a buyout option with a probability that is proportional to their number of products.

Labor Market Labor is split between production (L_p) and research labor. Research labor can be further subdivided into that devoted to private basic (L_b), public basic (L_u), private applied research (L_a) and firm entry (L_e). Since the total labor supply is M workers, the labor market clearing condition is given by

$$M = L_p + L_b + L_a + L_e + L_u.$$

The labor utilization from each component can be expressed in a more concise form when we investigate the properties of the dynamic equilibrium in the next section.

Household Problem Finally, we close the model by describing the household problem that determines the equilibrium interest rate in this model. The household consumes the final good and maximizes the following lifetime utility

$$W_0 = \int_0^\infty \exp(-\delta t) \frac{C(t)^{1-\gamma} - 1}{1-\gamma} dt \tag{10}$$

where $C(t)$ is consumption at time t, γ is the constant relative risk aversion parameter, and δ is the discount rate. The household owns all the firms in the economy, which generates a risk-free flow return of r in aggregate. The household also supplies labor in the economy, through which it earns wage rate $w(t)$. Finally, the household pays a lump-sum tax $T(t) \geq 0$ every instant. Thus, the household's intertemporal maximization is simply to maximize (10) subject to the following budget constraint

$$C(t) + \dot{A}(t) \leq r(t) A(t) + Mw(t) - T(t)$$

where $A(t)$ is the asset holdings of the household.

[13]The exact value of this price will not play any role for the equilibrium determination.

3.2 Equilibrium

In this section, we characterize the dynamic equilibrium of our model. Our focus is on a symmetric balanced-growth-path (SBGP) equilibrium where all industries start with the same initial conditions at time $t = 0$ and all aggregate variables grow at the same endogenous rate g.

In this model, three variables affect the payoff of the firm: the number of product lines n, the number of industries m, and the relative productivity

$$\hat{q}_{ij} \equiv q_{ij}/\bar{q}_i \tag{11}$$

of its product lines, which is the absolute productivity in line j normalized by the productivity index $\bar{q}_i$ in industry i. Thus, each incumbent firm is characterized by its state $k \equiv (\hat{\mathbf{q}}, n, m)$.

More specifically, given a government policy sequence $[T(t)]_{t=0}^{\infty}$, an SBGP equilibrium is composed of a sequence of intermediate good quantities, prices, the basic and applied innovation rates of private firms and entrants, the wage rate and interest rate, the joint distribution of multi-industry presence and product count, hot and cold product line productivity distributions, the fraction of hot product lines, i.e., $[y_k(t), p_k(t), b_k(t), a_k(t), a_e(t), w(t), r(t), \Gamma_{m,n}(t), \mathcal{F}_H(t), \mathcal{F}_L(t), \alpha(t)]_{t=0}^{\infty}$, such that all firms choose quantity and price to maximize their profits, incumbent and entrant firms invest in research to maximize their firm value, the labor market clears, the household maximizes its discounted sum of future utilities, and the distributions satisfy the relevant flow equations.

Solution of the Model The standard monopoly profit maximization delivers the following familiar equilibrium price and quantities (interested readers are referred to the online appendix Section OA-1 for the detailed derivations)

$$y_j = \hat{q}_j^{\epsilon} Z \qquad \text{and} \qquad p_j = \frac{1}{M\hat{q}_j}. \tag{12}$$

Clearly, a monopolist's quantity is increasing and price decreasing in the relative productivity $\hat{q}$ of the product line. Finally, the equilibrium profits of the monopolist are again increasing in its relative productivity $\hat{q}$ and the average market size Z/M:

$$\pi(\hat{q}) = \frac{\hat{q}^{\varepsilon-1}}{\varepsilon}\frac{Z}{M}. \tag{13}$$

Next, only in this section, we focus on myopic firms that maximize their one-period-ahead returns (as opposed to forward-looking firms that maximize the discounted sum of future profits). This will allow us to provide some useful analytical results and highlight the key economic forces of our model. In our quantitative analysis (Section 5), we will generalize this and focus on forward-looking firms.

Myopic Firms Consider now a firm that has n product lines in m industries. Moreover, in an SBGP, an α fraction of product lines are hot. Then the maximization problem when deciding for the amount of basic research can be written as

$$\max_{b_m} \left\{ n b_m \left(1+\rho_m\right) V^H - \tilde{w} n b^{\nu_b} \xi_b \right\}$$

where $V^H \equiv \mathbb{E}^H_{\hat{q}} \pi\left(\hat{q}+\eta\right)$ is the expected return to a *successful* basic innovation and $\tilde{w} \equiv \frac{w}{Z/M}$ is the normalized wage rate. Several observations are in order. First, the expected return from basic research investment is increasing as the firm has fingers in more pies as Nelson argued (higher ρ_m). Second, the innovations are undirected within industries; therefore, the firm has to form an expectation for the expected profit $\mathbb{E}^H_{\hat{q}} \pi\left(\hat{q}+\eta\right)$, which means that we have to keep track of the invariant relative productivity distribution to compute V^H. Finally, both the returns and the costs are proportional to the number of product lines n, which makes the problem much more tractable and the quantitative solution manageable. Now we can express the first-order condition as

$$b_m = \left[\frac{\left(1+\rho_m\right) V^H}{\nu_b \xi_b \tilde{w}} \right]^{\frac{1}{\nu_b - 1}}$$

The most important result here is the fact that basic research investment is increasing in the multi-industry presence of the firm. The strength of this positive relationship will be mainly governed by the probability of the cross-industry spillover parameter p, which will help us match Figure 4.

Fact 1 *A firm's basic research investment is increasing in its multi-industry presence.*

Both private firms and public research labs are generating basic research in this economy. It is useful to break down total basic research into its embodied and disembodied components. The distinction is based on whether the basic knowledge is immediately turned into a consumer product (embodied) or simply added to the stock of knowledge available for future innovators (disembodied). We obtain the following aggregates

$$\begin{aligned} \text{Embodied:} \quad \tau_b^e &\equiv \sum_{m=1}^{M} \mu_m (1+\rho_m) b_m \\ \text{Disembodied:} \quad \tau_b^d &\equiv \sum_{m=1}^{M} \mu_m (p-\rho_m) b_m + (1+p) u \\ \text{Total:} \quad \tau_b &\equiv \tau_b^e + \tau_b^d \end{aligned} \tag{14}$$

where we define the mass of product lines owned by firms in m industries by μ_m, which can be computed from the joint distribution using $\mu_m \equiv \sum_{n=1}^{\infty} n \cdot \Gamma_{m,n}$. Then τ_b^e and τ_b^d correspond respectively to the embodied and disembodied components of basic research. Note that the disembodied component includes both private spillovers that are unused and the results of public basic innovation. Finally, τ_b is simply the overall flow of basic innovation, including all spillovers.

Using this aggregate rate and the cool-down rate ζ, we can express the steady-state flow equation: the number of product lines that become hot must be equal to the number of product lines that cool down. In other words, we must have $\alpha\zeta = (1-\alpha)\tau_b$. As a result, the steady-state fraction of hot product lines is

$$\alpha = \frac{\tau_b}{\zeta + \tau_b}. \tag{15}$$

The share of hot product lines, those having basic knowledge that can be turned into better consumer products (α), is increasing in the amount of basic research flow. This expression highlights the role of public policy in affecting the knowledge stock. The more money is allocated to public basic research, the higher will be the basic research flow from public research labs (u), which will then increase the fraction of hot product lines through τ_b, as in (14) and (15).

However, a bigger α is meaningful only when there is subsequent applied research that turns this existing basic knowledge stock into consumer products. Therefore, we now turn to the applied research decision of the firms. Their maximization problem is simply

$$\max_a \left\{ na \left[\alpha V^H + (1-\alpha) V^C \right] - \tilde{w} n a^{\nu_a} \xi_a \right\}$$

where $V^H \equiv \mathbb{E}^H_{\hat{q}} \pi \left(\hat{q} + \eta \right)$ is the expected returns from hot product lines and $V^C \equiv \mathbb{E}^C_{\hat{q}} \pi \left(\hat{q} + \lambda \right)$ is that from cold ones and $\tilde{w} = \frac{w}{Z/M}$ is the normalized wage rate. When investing in applied research, firms form two types of expectations. The first one is due to the undirected nature of research: firms have to form expectations over the relative productivity $\hat{q}$ that they are going to land on. The second, and more important one, is due to the complementarity between basic and applied research: firms take into account the fraction of hot product lines. Firms invest in applied research according to

$$a = \left[\frac{\alpha V^H + (1-\alpha) V^C}{\nu_a \xi_a \tilde{w}} \right]^{\frac{1}{\nu_a - 1}}.$$

The crucial observation here is the complementarity between basic and applied research. In equilibrium $V^H > V^C$ since hot product lines are associated with a larger step size η. Hence, if there are more hot product lines (a higher α), each firm increases its investment in applied research.

Fact 2 *Basic and applied research investments are complementary. In particular, higher public basic research investment encourages firms to invest more in applied research.*

However, the fraction of hot product lines α is not sufficient to determine the incentives for applied research alone due to the correlation between this product state and productivity. The incentives will be a function of the fraction of hot and cold product lines and the average qualities within those types. In particular, firms must know the values of $\mathbb{E}^H_{\hat{q}} (\hat{q} + \eta)^{\varepsilon - 1}$ and $\mathbb{E}^C_{\hat{q}} (\hat{q} + \lambda)^{\varepsilon - 1}$ due to the exact form of the profit function in equation (13). Therefore, Lemma 1 describes the laws of motion for the type-specific productivity distributions.

Let us denote the aggregate rate of applied innovation by τ_a such that

$$\tau_a = \sum_{m=1}^{M} \mu_m a_m + E a_e. \tag{16}$$

Note that in the baseline model, $a_m = a$ for all m, but this will not necessarily be the case in the general model in Section 4. Recall that τ_b^e denotes the arrival rate of embodied basic research, as defined in (14). Now we can denote the aggregate rate of *creative destruction* (the rate at which firms lose product lines to other firms) by τ:

$$\tau \equiv \tau_a + \tau_b^e. \tag{17}$$

Creative destruction is determined by the rate at which incumbents produce basic innovations which can be embodied into production immediately (τ_b^e), and by the rate at which incumbents and entrants produce applied innovations (τ_a). Now we are ready to state the following lemma.

Lemma 1 *Let $\mathcal{F}_H(\cdot,t)$ and $\mathcal{F}_C(\cdot,t)$ be the aggregate product cumulative measures by type (hot or cold). The flow equations for these objects are, respectively,*

$$\dot{\mathcal{F}}_H(\hat{q}) = -\tau\left[\mathcal{F}_H(\hat{q}) - \mathcal{F}_H(\hat{q}-\eta)\right] + \tau_b^e \mathcal{F}_C(\hat{q}-\eta) - \zeta\mathcal{F}_H(\hat{q}) + \tau_b^d \mathcal{F}_C(\hat{q}) + g\hat{q}[\partial\mathcal{F}_H(\hat{q})/\partial\hat{q}]$$
$$\dot{\mathcal{F}}_C(\hat{q}) = -\tau_a\left[\mathcal{F}_C(\hat{q}) - \mathcal{F}_C(\hat{q}-\lambda)\right] - \tau_b \mathcal{F}_C(\hat{q}) + \zeta\mathcal{F}_H(\hat{q}) + g\hat{q}[\partial\mathcal{F}_C(\hat{q})/\partial\hat{q}]$$

Proof. See Appendix OA-1. ■

The labor market clearing condition can now be expressed in terms of the above endogenous variables. One additional relationship we will exploit is that between the mass of labor devoted to production and the normalized wage rate. This can be derived from the goods production specification (see Section OA-1 in the online appendix for its detailed derivation)

$$L_p = \frac{Z}{w}\left(\frac{\varepsilon-1}{\varepsilon}\right)$$

Using this and the symmetric nature of the equilibrium, we express the labor market clearing condition as an average over industries

$$1 = \frac{1}{\tilde{w}}\left(\frac{\varepsilon-1}{\varepsilon}\right) + \xi_b\left(\sum_m \mu_m b_m^{\nu_b} + U u^{\nu_b}\right) + \xi_a(a^{\nu_a} + E a_e^{\nu_a}) \tag{18}$$

This expression equates the labor supply per industry (= 1 since the total labor supply is M) to labor demand for production workers; private basic researchers, which is a function of the multi-industry presence of the firms; public basic researchers, which is determined by public policy; incumbent applied researchers; and entrant basic researchers.

Finally, plugging the equilibrium intermediate good quantity (12) into the aggregate production functions (2) and (1), we find that the aggregate output is

$$Z = \bar{q} L_p / M \tag{19}$$

This expression simply says that the aggregate output is equal to the product of the number of workers employed for production and the aggregate productivity index of the economy. In an SBGP equilibrium, the labor allocated for production is constant. Therefore the growth rate of aggregate output (and also output per worker) will be equal to the growth rate of the productivity index $\bar{q}$. The following proposition provides the exact growth rate of the productivity index.

Proposition 1 *In an SBGP, the growth rate of the productivity index is*

$$g = \frac{\tau_a \left[\alpha \mathbb{E}_{\hat{q}}^{H} \left(\hat{q}+\eta\right)^{\varepsilon-1} + (1-\alpha)\mathbb{E}_{\hat{q}}^{C} \left(\hat{q}+\lambda\right)^{\varepsilon-1} - 1\right] + \tau_b^e \left[\mathbb{E}_{\hat{q}} \left(\hat{q}+\eta\right)^{\varepsilon-1} - 1\right]}{\varepsilon - 1} \tag{20}$$

Proof. See Appendix A ■

This growth expression shows that the engines of economic progress include both applied and basic innovation. More important, the basic knowledge stock in the economy, represented by α, makes each applied innovation more valuable and contributes more to growth (since $\eta > \lambda$). This expression shows how public funding can contribute to growth through its indirect impact on private research.

3.3 Discussion of the Model

In this section, we briefly discuss sources of inefficiency and what policy can achieve in this model. First, as in standard quality ladder models, there are intertemporal spillovers within each product line. Second, firms simply enjoy the expected duration of monopoly power due to the competition channel of creative destruction. As a result, the private value of innovation differs from the social value of innovation. It is also worth highlighting that in this model, there could be either over or underinvestment in R&D. In addition to the standard channels, our model features additional spillovers due to basic research, both within and across industries. Finally, there are additional static distortions due to monopoly power. However, since we are primarily interested in the dynamic inefficiencies associated with innovation and basic research, we will consider the case of a social planner who is still subject to monopoly distortions on the production side.

All of these inefficiencies will generate room for innovation policy, and our estimated model will govern whether there is over- or underinvestment in the various types of research expenditures in the decentralized equilibrium. It will also provide a framework within which to evaluate the effects of these innovation policies.

4 Generalizations of the Model

This section generalizes the baseline model to provide richer dynamics for the economy and its agents. Those not interested in the technical details can skip directly to the quantitative Section 5.

Stochastic Innovation Step Sizes In the general model, as in Klette and Kortum (2004) and Lentz and Mortensen (2008), we assume that innovation step sizes are drawn from exponential distributions. For basic research, the mean of the distribution is always η. For applied research, the distribution mean is η if the product line is hot and λ if it is cold. As in the baseline model, $\eta > \lambda$.

Fixed Cost of Basic Research In our sample, some firms do not invest in basic research. To capture this fact, we generalize the basic research technology by introducing a fixed cost of doing basic research. At each instant, a firm with n product lines draws a fixed labor cost of doing basic research $n\phi \geq 0$, where ϕ is distributed according to the distribution $\mathcal{B}(\cdot)$. Then a firm that operates in n product lines and has a fixed cost of basic research ϕ this period has the following cost function $C_b(b_m \mid n, \phi) = nc_b(b_m \mid \phi)$. This implies that firms will follow a cutoff rule as a function of their multi-industry presence ϕ^*_m such that they will not invest in basic research if $\phi > \phi^*_m$. Otherwise, in addition to the variable cost, they will also pay the fixed cost.

Industry Expansion (Start-up Buy-outs) In the baseline model, we took the working knowledge of the firms (m) as exogenously given. We now endogenize m by introducing the possibility of buy-out offers for new entrants. The economy features $E \times a_e$ flow of entry at any instant. We will assume that a ς fraction of new entrants will meet a randomly selected incumbent firm. Thus, an incumbent will have a flow rate of incoming buy-out offers

$$x \equiv \varsigma E a_e / F.$$

where F is the equilibrium measure of firms. If $\bar{n}$ denotes the average number of product lines per firm, then $F = 1/\bar{n}$. Clearly this new company will be from a new industry with probability $(1 - m/M)$ or from an industry that already exists in the incumbent's portfolio with probability m/M. Our goal is to keep the M&A margin as tractable as possible, and we will achieve this by assuming that the M&A price that the incumbent firm has to pay is equal to the full surplus of the new merger. The resulting invariant joint distribution $\Gamma_{m,n}$ over multi-industry presence m and firm product count n is described in Appendix A.

Forward-Looking Firms For expositional purposes, in the previous section we described the model with myopic firms that maximize their one-period-ahead returns. For the rest of our analysis, we relax this assumption and consider firms that maximize the discounted sum of future returns. The analysis of this new model is very similar to that of the previous model except that the returns to innovation take the form of a value function that takes into account all future contingencies. The following proposition provides the exact forms of the value of a firm that has a productivity portfolio $\hat{\mathbf{q}}$ and operates in m industries.

Proposition 2 *Let the value of a firm with a productivity portfolio* $\hat{\mathbf{q}}$ *in* m *industries be denoted by* $\mathcal{V}(\hat{\mathbf{q}}, m)$*. This value is equal to*

$$\mathcal{V}(\hat{\mathbf{q}}, m) = \frac{Z}{M}\left[\sum_{\hat{q}\in\hat{\mathbf{q}}} V(\hat{q}) + nV_m\right]$$

where

$$V(\hat{q}) = \frac{\hat{q}^{\epsilon-1}}{\varepsilon\left[r+\tau+\kappa+g(\epsilon-2)\right]}$$

and

$$(r-g)V_m = \max_{a,b}\left\{\begin{array}{c} -\tilde{w}\left[h_a(a)+h_b(b)+\mathbf{1}_{(b>0)}\phi\right] \\ +a\left[\alpha V^H+(1-\alpha)V^C+V_m\right]+b(1+\rho_m)\left[V^H+V_m\right] \\ +x\left(1-\frac{m}{M}\right)\left[V_{m+1}-V_m\right]-\tau V_m+\kappa\mathbb{E}_{\hat{q}}V(\hat{q}_t)\end{array}\right\}. \tag{21}$$

The analogous production values are defined as $V^H \equiv \mathbb{E}^H_{\hat{q},\eta}V(\hat{q}+\eta)$ *and* $V^C \equiv \mathbb{E}^C_{\hat{q},\lambda}V(\hat{q}+\lambda)$*.*

Proof. See Appendix A ■

This important result has a number of implications. First, the value of a firm has a tractable additive form across product lines. Moreover, the firm value has two major components: the first component is the production value $V(\hat{q})$, which simply computes the sum of the future discounted profits where the effective discount rate takes into account the rate of creative destruction τ, the exogenous destruction rate κ, and the obsolescence of the relative productivity $\hat{q}$ due to the growth of $\bar{q}$. The second component is the R&D option value V_m, which is a direct function of the multi-industry presence due to the associated internalization of spillovers. Finally, because of the stochastic nature of step sizes, the expectations now integrate over the productivity (which are type specific) and step size.

Welfare Finally, we close this section by describing the SBGP equilibrium welfare. In an SBGP equilibrium that has an initial consumption C_0 and a growth rate of g, welfare is computed as

$$W(C_0,g)^{SBGP} = \int_0^\infty \exp(-\delta t)\frac{\left(C_0 e^{gt}\right)^{1-\gamma}}{1-\gamma}dt = \frac{1}{1-\gamma}\left(\frac{C_0^{\frac{1-\gamma}{\varepsilon-1}}}{\rho-(1-\gamma)g}-\frac{1}{\rho}\right)$$

We will report our results in consumption-equivalent terms. In particular, when two different public policies T_1 and T_2 generate different SBGP equilibrium welfare values as $W(C_0^{T_1}, g^{T_1})$ and $W(C_0^{T_2}, g^{T_2})$, we will report β such that

$$W(\beta C_0^{T_1}, g^{T_1}) = W(C_0^{T_2}, g^{T_2}).$$

In other words, β constitutes the compensating differential in initial consumption that equalizes the welfare of the two proposed policy environments. It therefore provides an intuitive measure for evaluating policy tools. This completes the description of the theoretical environment. Now we are ready to move on to the quantitative analysis.

5 Quantitative Analysis

In this section we describe the estimation strategy used. We will assume that the fixed costs are drawn from a lognormal distribution $\mathcal{B}(\phi)$ with mean $\bar{\phi}$ and variance σ^2. As a result, the set of parameters of the model is

$$\theta = \{\delta, \gamma, \varepsilon, p, \eta, \lambda, E, U, \zeta, \nu_a, \nu_b, \xi_a, \xi_b, \kappa, \bar{\phi}, \sigma\} \in \Theta.$$

During the period we consider, there was existing government support for R&D activities in France. Our data set contains information on the publicly funded portion of private R&D. On average, 10% of private R&D was funded publicly. Therefore in our estimation, we introduce a uniform subsidy to the total R&D spending of the firm $\psi = 0.10$. The government has a balanced budget every period, so that the sum of total subsidies (S) and public research funding (R) must be equal to tax revenues, that is

$$T = S + R = \psi \left[\sum_{m=1}^{M} \mu_m C_B(b_m \mid \phi) + C_A(a) \right] + U C_B(u \mid \bar{\phi})$$

where T is a lump-sum tax on consumers. In France, during 2000-2006, the fraction of GDP devoted to public research labs and academic institutions was approximately 0.5%. Therefore, we pick R/Z, which is the share of GDP devoted to public basic research, to be 0.5%.

5.1 Estimation Method

In our data set, for each firm f and each time period t, we have a vector of N observables from the actual data $\mathbf{y}_{ft} \equiv \left[y_{ft}^1 \quad . \quad . \quad . \quad y_{ft}^N \right]'_{N \times 1}$ including the number of industries in which the firm is present, sales, profits, and labor costs. Let the entire data set be denoted by $\mathbf{y}$.

We use the simulated method of moments (SMM) for the estimation.[14] Define $\Lambda(\mathbf{y})$ and $\Lambda(\theta)$ to be, respectively, the vectors of real data moments (generated from $\mathbf{y}$) and equilibrium model moments (generated for some vector of parameters θ). Since certain moments require knowledge of the joint distribution of firms over the number of products and industries (m, n) and the portfolio of product qualities $\mathbf{q}$, which has no apparent analytic form, we simulate a large panel of firms to calculate $\Lambda(\theta)$ to a high degree of accuracy.[15]

Our proposed estimator minimizes a quadratic form of the difference between these two vectors

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} \left[\Lambda(\theta) - \Lambda(\mathbf{y})\right] \cdot W \cdot \left[\Lambda(\theta) - \Lambda(\mathbf{y})\right]$$

where W is the weighting matrix. We use a diagonal weighting matrix with entries equal to the inverse square of the data moment value, or in notational terms $W_{ii} = 1/\Lambda_i(\mathbf{y})^2$ and $W_{ij} = 0$

[14] See Bloom (2009) and Lentz and Mortensen (2008) for further description and usage information on SMM.

[15] For our results, we simulate 32K firms with a burn-in time of 100 years and 100 time steps per year.

for $i \neq j$. In our estimation, we use 26 moments. We pick moments that are most informative for the unique features of our model. In particular, we target both the intensive and extensive margins of basic research intensity as it varies with multi-industry presence. Since multi-industry presence is one of the key determinants of innovation, we target both the mean and the variance of that quantity. In addition, we include aggregate and conditional firm-level growth rates. Since our model is macro growth model and household's welfare (and accordingly the policy analysis) depends crucially on the level of aggregate growth, hitting that moment is of particular importance. For that purpose, we boost the weighting on the aggregate growth moment.[16] To capture the within-industry spillover, we target the spillover differentials described in Section 2.4. Finally, to further inform the model parameters on firm dynamics, we include the mean return on sales, the R&D/production labor ratio, the exit rate, and mean firm age by size. The details of the moments and identification are described in Online Appendix Section OA-4.

5.2 Computer Algorithm Outline

An equilibrium of this model is described by a system of five equations in the five variables $(\tau_a, \tau_b^e, \tau_b^d, \tilde{w}, g)$. This system can be evaluated using the following procedure:

1. Calculate α and the distribution over $\hat{q}$ using τ_a, τ_b^e, τ_b^d, and g according to equations (1) and (15).

2. Calculate g using, τ_a, τ_b, and the distribution over $\hat{q}$ with equation (20).

4. Calculate $V^H = \mathbb{E}^H_{\hat{q},\eta} V(\hat{q}+\eta)$ and $V^C = \mathbb{E}^C_{\hat{q},\lambda} V(\hat{q}+\lambda)$ using the relevant step size distribution and the type-specific productivity distributions.

5. Find a_m and b_m using first-order conditions with $\tilde{w}$ from equation (21).

6. Impose an upper bound on n and find the steady state $\Gamma_{m,n}$ using the flow equations in A.

7. Compute the updated values of τ_a, τ_b^e, and τ_b^d using (16) and (14).

8. The difference between the conjectured and updated values of τ_a, τ_b^e, τ_b^d, and g in conjunction with the labor market clearing differential from (18) constitute the five desired equations.

We use Powell's (Powell (1970)) hybrid equation solver to solve this set of equations for a given set of parameters. To minimize the SMM objective function, we perform a search over the parameter space using a combination of a naive simulated annealing algorithm and a Nelder-Mead simplex (Nelder and Mead (1965)) algorithm. See Zangwill and Garcia (1981) for more information on solving systems of nonlinear equations.

Table 5: Parameter Estimates

#	Description	Sym	Value	#	Description	Sym	Value
1.	Discount Rate	δ	0.038	9.	Applied Cost Curvature	ν_a	1.367
2.	CRRA Utility Parameter	γ	2.933	10.	Basic Cost Curvature	ν_b	1.538
3.	Elasticity of Substitution	ε	5.800	11.	Applied Cost Scale	ξ_a	1.228
4.	Cross-industry Spillover	p	0.113	12.	Basic Cost Scale	ξ_b	5.437
5.	Basic Step Size	η	0.079	13.	Exogenous Exit Rate	κ	0.006
6.	Applied Step Size	λ	0.049	14.	Basic Fixed Mean	$\bar{\phi}$	-4.761
7.	Mass of Entrants	E	0.502	15.	Basic Fixed Std. Dev.	σ	0.327
8.	Mass of Academic Labs	U	0.491	16.	Product Cooldown Rate	ζ	0.116

5.3 Estimation Results

Table 5 reports the values of the estimated structural parameters. The estimated values of the discount rate and CRRA utility parameters are within their standard macro ranges. The elasticity of substitution parameter generates 17%$(= 1/\varepsilon)$ gross profits, resulting in 7.9% net profits after subtracting R&D expenses as a share of sales.

One of the most important parameters of our model is the cross-industry spillover parameter $p = 0.11$, which measures the probability that a basic innovation will have an additional immediate application. This estimate affects the extent to which basic innovations contribute to cross-sectional growth. In equilibrium, firms operate in two industries out of 10 on average. Therefore, any given spillover is not embodied with probability 89%$(= 8/9)$. Given that the probability of having a spillover is 11%, the probability of having a disembodied spillover is 10%$(= 0.11 * 0.89)$.

The estimated innovation size of basic research is $\eta = 7.9\%$ and the innovation size of each new applied innovation is $\lambda = 4.9\%$. This implies that basic research (hot product lines) makes applied innovation 60%$(= 7.9/4.9 - 1)$ more productive.

Additionally, each basic innovation has a within-industry spillover. The cool-down rate of hot product lines is estimated to be $\zeta = 0.12$, which indicates that a basic innovation affects the subsequent innovations in the same product line for almost 8.3$(= 1/0.12)$ years on average.

The elasticity of applied innovation counts with respect to the research dollars spent is estimated to be 0.73 $(= 1/\nu_a)$ and similarly the elasticity of basic innovation with respect to the basic research investment is 0.65 $(= 1/\nu_b)$. These values are close to the elasticity estimates in the literature, which typically finds a value around 0.5 (Blundell, Griffith, and Windmeijer (2002), Griliches (1990), Pakes and Griliches (1984) and Kortum (1992, 1993)).

5.4 Goodness of Fit

In this section, we will first focus on the moments that we targeted in our estimation and then turn to the moments that we did not directly target but still find useful in understanding the model's

[16] Increasing the weighthing factor to 3 was sufficient to align the aggregate growth rate in the data and the model.

performance.

Targeted Moments Table 6 contains the values of moments from the actual data and our estimated model.

Table 6: Moments Used in Estimation

#	Description	Model	Data	#	Description	Model	Data
1-8	Basic Research Extensive	See Figure 10		21	R&D/Labor	0.284	0.260
9-16	Basic Research Intensive	See Figure 9		22	Employment Growth	0.111	0.103
17	Mean Industries	2.217	2.203	23	Aggregate Growth	0.013	0.015
18	Mean Square Industries	7.213	6.975	24	Spillover Differential	8.378	8.000
19	Return on Sales	0.032	0.032	25	Age, Small Firms	11.53	14.99
20	Exit Rate	0.082	0.091	26	Age, Large Firms	18.69	24.87

The results indicate that the model performs very well in generating firm and industry dynamics similar to those in the data. As documented in Section 2.1, a significant fraction of innovating firms invest in basic research. In particular, 29% of firms are investing in basic research, which was 27% in the data. We also capture the positive relationship between the extensive margin of basic research and multi-industry presence, as evidenced in Table 6 and Figure 9.

Figure 9: Fraction Positive Basic By # Industries

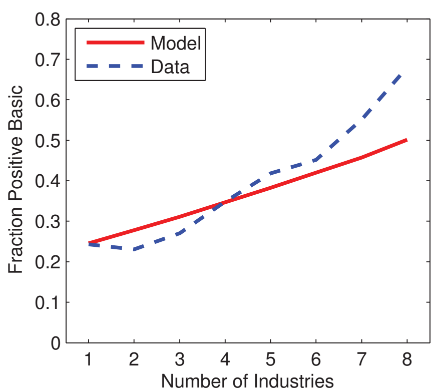
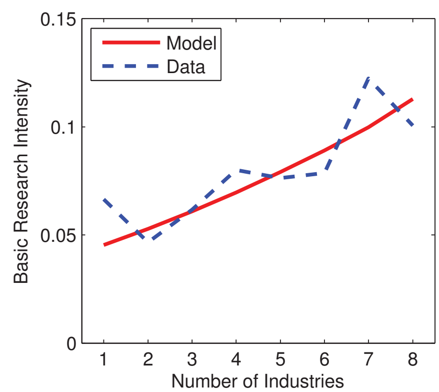


Figure 10: Basic Research Intensity By # Industries



The positive correlation between multi-industry presence of a firm and its basic research intensity was one of the primary motivations for introducing multi-industry presence into our model. As explained previously in the text, multi-industry presence plays an important role in increasing basic research incentives, by allowing a greater potential to internalize the positive spillovers from basic research. In our reduced-form analysis, we found a significant and positive correlation between multi-industry presence and basic research intensity. This has been the key moment to identify the cross-industry spillover parameter. Our model successfully generates this positive correlation.

In the data, firms operate on average in 2.2 industries, and the same is true in the model.

Furthermore, we find the mean squared in the model to be 7.2, compared to 7.0 observed in the data.

The table above reports some additional moments that are not captured by the stylized facts. For instance, the mean profitability is 3.2% in the model and in the data. The prime determinants of profitability are the step sizes for basic and applied innovation. However, these also affect the investment levels for both types of research, since this increases the return to successful innovation. Therefore, the step size parameters are set to be a compromise between hitting the profitability moment and the research investment and growth moments.

We are targeting additional moments regarding research investments. The first is the average ratio of total research labor to production labor by incumbent firms. The model comes very close to hitting this ratio exactly (28.4% vs 26.0%), largely in order to hit the aggregate growth and return on sales.

All of these components of the economy determine the aggregate growth rate. Our model matches the observed growth rate closely. Our model economy grows at a rate of 1.3%, while the French economy grew at an average rate of 1.5% during the period studied (2000-2006).

Untargeted Moments In this part, we discuss our model's prediction about some of the moments that we did not directly target.

Interestingly in the data the correlation between profitability and basic research intensity is not significantly different from zero. The same implication emerges from our model. In the baseline model, the correlation between profitability and basic research intensity is only 0.04. This result emerges because basic research investment is determined through the multi-industry presence of the firms, whereas profitability is determined by the share of hot and cold product lines, type of research investment, and the productivity distribution $\mathcal{F}(\hat{q})$ in the economy.

Our model naturally generates a positive correlation between multi-industry presence and firm size, which is also empirically true in the data. This arises since both of these moments are strongly correlated with firm survival. In the model, we find a correlation of 0.52 between the log employment and multi-industry presence. In the data, this value is 0.76.

Another stylized fact in our data is that the firm size distribution is highly skewed. This is a well-known feature that is documented extensively in the literature. For detailed references, see Aghion, Akcigit, and Howitt (2013). In our model, we capture this fact with a skewness of the firm size distribution of 4.12. This value is 3.07 in the data.

Our estimates indicate that entrants play an important direct role in overall growth. The innovation rate from entrants is 0.43%, whereas that number is 0.92% for incumbents. That implies that entrants account for 32% of growth. Though our number is for the French economy, our number is in line with Foster, Haltiwanger, and Krizan (2001) who find that 25% of productivity growth in the US comes from new entry.

We will now focus on the details of the equilibrium and the social planner's problem to study the efficiency properties of this economy. Then we will turn to various policies that could address this inefficiency.

5.5 Endogenous Variables of the Baseline Economy

The following table provides equilibrium values for some of the important endogenous variables in the model:

Table 7: Decentralized Economy: Endogenous Variables (in percentages)

ψ	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
10	22.0	0.58	0.28	85.6	0.53	0.52	4.5	8.9	6.9	1.34	95.3

In this table, τ_a denotes the aggregate rate of applied innovation by incumbents and entrants, whereas τ_b^e and τ_b^d denote the aggregate rates of embodied and disembodied basic innovation, respectively. The next five columns report the labor allocations into production, private basic, public basic, entry, and applied research. The remaining columns report the fraction of hot product lines α, the ratio of consumption to that for the social planner's economy, the growth rate g, and the welfare in consumption equivalent terms β.

The model highlights the dynamic misallocation of research effort and its welfare consequences. In our benchmark economy, 85.6% of labor is used for production, and 14.4% is employed for innovation activities. Among researchers, roughly 7% are engaged in basic research activities. Note that this composition within innovation activities will be the central focus of the policy analysis, since uninternalized spillovers are one of the main sources of inefficiency. As a consequence, the arrival rate of basic innovation in our baseline economy is significantly smaller (25 times) than that of applied innovation. This translates into significant welfare losses, with the economy achieving only 95.3% of welfare with respect to the social planner's optimum, which we will analyze next.

5.6 Quantifying the Social Planner's Optimum

In this section, we are going to provide the solution to the social planner's problem under two scenarios. In the first, the planner, as in the Ivory Tower approach of the baseline case, cannot appropriate public basic research returns. This is illustrated in Panel A. In the second, the planner can appropriate and use new basic inventions for production immediately. This case is reported in Panel B. We will set the welfare to 100% in this case and report the remaining welfare numbers relative to this baseline. Finally, recall that we are considering a planner who controls firm's research labs but not their production decisions. The following table summarizes these results.

One striking feature of the solution to the social planner's problem under both scenarios is that the fraction of labor devoted to research activities is not substantially greater than in the

Table 8: Social Planner's Optimum (in percentages)

τ_a^{SP}	$\tau_b^{e,SP}$	$\tau_b^{d,SP}$	L_p^{SP}	L_b^{SP}	L_u^{SP}	L_e^{SP}	L_a^{SP}	α	g	β
A. Non-appropriated Public Research										
19.1	5.1	0.2	82.9	5.6	0.5	3.7	7.3	31.1	1.80	98.7
B. Appropriated Public Research										
18.5	6.5	0.0	82.6	4.6	2.3	3.5	7.0	35.8	1.93	100

decentralized equilibrium. In particular, in Panel A the total labor allocated to research activities was 14% in the decentralized economy, while it is only 17% when set by the social planner.

Indeed, the dominant misallocation here is not that between production and research, as is common in this class of models, but among the various types of research activities, in this case, applied and basic innovation. In the decentralized economy, only 1.05% of the total labor force is devoted to basic research, whereas in the social planner's economy, this number rises to 6.1%. In other words, the social planner devotes 36% of research labor to basic research, whereas this fraction was only 7% in the decentralized economy. This happens on both the intensive and the extensive margins of basic research. In fact, the planner finds it optimal to employ nearly all private research labs, regardless of their fixed cost draw.

Another interesting and important finding is that in the case of applied innovation, there is actually an *overinvestment* in the baseline economy. The applied research labor utilization (including entrants) is 13.6% in the decentralized case. This figure drops to 11% in the social planner's solution. This is in spite of the fact that the fraction of hot product lines rises from 7% to 31%, meaning the average step size of an applied innovation rises by almost a third.

The net result of the above changes is that growth rises to 1.8% from 1.34%. Overall, the decentralized economy's welfare corresponds to a decrease of 3.4%($= 1-95.3/98.7$) in consumption-equivalent terms from the social planner's optimum. The following policy experiments will try to bridge this gap.

Panel B reports these numbers for the case of appropriated public research. The main difference is the sizable increase in the labor devoted to public basic research, which rises to 2.3% relative to 0.5% in both the decentralized economy and Panel A. When public basic research turns into production immediately, this contributes to aggregate growth by an additional 0.13 percentage point and increases welfare by an additional 1.3 percentage points in consumption equivalent terms. Policies such as the Bayh-Dole Act allow academic researchers to appropriate their innovations through patenting. In our setting, this would correspond to an increase in the rate of appropriation of innovation by public researchers. We will consider this as a policy tool in Section 6.3.

6 Policy Analysis

In this section, we analyze the impact of different types of research subsidies. Given our distinction between basic and applied research, it seems natural to propose different subsidy policies for different types of research spending. However, this could potentially generate a moral hazard problem, since firms would have an incentive to misreport the type of research they undertake, which is very difficult for a policymaker to verify. However, it is still useful to consider this hypothetical case to form a benchmark.

This section is organized as follows: Section 6.1 starts with this hypothetical case, Section 6.2 considers a uniform research subsidy as in the real world, Section 6.3 considers only optimal funding of public research labs, and finally Section 6.4 combines both uniform subsidy and public research funding using feasible policy tools.

6.1 Type-Dependent Research Subsidy

Assume first that the policymaker can distinguish between different types of research efforts and accordingly provide differentiated subsidy rates. Let ψ_a and ψ_b denote the applied research and basic research subsidy rates, respectively. The share of GDP allocated to public research (R/Z) is kept constant by the policymaker. Note that an increase in the subsidy rate (ψ_a or ψ_b) reduces research costs for the firm and leads to an increase in research effort as a result. The following table reports the optimal subsidy rates and resulting equilibrium variables.

Table 9: Type-Dependent Research Subsidy (in percentages)

ψ_a^{TD}	ψ_b^{TD}	τ_a^{TD}	$\tau_b^{e,TD}$	$\tau_b^{d,TD}$	L_p^{TD}	L_b^{TD}	L_u^{TD}	L_e^{TD}	L_a^{TD}	α^{TD}	g^{TD}	β^{TD}
14	50	19.3	4.50	0.38	83.1	5.3	0.50	3.7	7.5	29.6	1.75	98.2

Since the underinvestment is mainly in basic research, the optimal type-dependent subsidy dictates a much larger subsidy rate for it, namely, $\psi_b = 50\%$ and $\psi_a = 14\%$. Here, the major component of policy is a fivefold increase in the subsidy rate for basic research, whereas the subsidy rate on applied innovation remains roughly the same.

The large value for the basic research subsidy is straightforward to understand. There are spillovers associated with basic innovation that are not internalized by firms. By subsidizing this type of innovation, we can mitigate this effect. This policy can almost achieve the level of welfare seen in the relevant social planner's case in Panel A of Table 8 (98.2% vs 98.7%).

As discussed above, this policy is hard to implement in the real world due to the moral hazard problem. Therefore, we focus on a policy providing a uniform subsidy across the economy.

6.2 Uniform Private Research Subsidy

With this policy, the government subsidizes a fraction ψ of each firm's total research investment, keeping the share of funds allocated to academic research constant. Note that such a policy subsidizes not only basic research but also applied research. The following table summarizes the results of the optimal uniform subsidy rate.

Table 10: Uniform Research Subsidy (in percentages)

ψ^{UP}	τ_a^{UP}	$\tau_b^{e,UP}$	$\tau_b^{d,UP}$	L_p^{UP}	L_b^{UP}	L_u^{UP}	L_e^{UP}	L_a^{UP}	α^{UP}	g^{UP}	β^{UP}
31	25.4	1.52	0.26	81.8	1.54	0.49	5.41	10.8	13.2	1.70	96.1

Our analysis of the baseline economy and the planner's economy documented a slight underinvestment in research overall and a large misallocation between the different types of research. A uniform subsidy is therefore ill suited to address these issues as it cannot directly affect the allocation between research types, and any attempt to subsidize basic research will only worsen the overinvestment in applied research. Although the optimal type-dependent basic subsidy is 50%, the optimal uniform subsidy is only 31%, due to cross-subsidization of applied research whose optimal level was 14%.

Under this policy, we are allocating a larger fraction of the labor force to research relative to the social planner's economy. Overall, the researcher's share goes up to 18% from 14%. As a result, we have too few workers devoted to production of the consumption good (81.8%) relative to the social planner's allocation (82.9%), which reduces the initial consumption of the baseline economy. Even though we have more labor working for research, the economy grows at a lower rate (1.7%) than the social planner's (1.8%). This interesting result emerges due to the misallocation of researchers between basic and applied innovation. The welfare gain from this policy is 0.8 percentage points, which is only 28%($= 0.8/2.9$) as large as that for the type-dependent policy.

Although the underinvestment in basic research is sizable, the uniform policy partially makes up for this at the cost of worsening the overinvestment in applied research. The main lesson to be drawn from this is that a uniform research subsidy should take into account its negative welfare consequences through its oversubsidization of applied research. Finding a feasible method to differentiate basic and applied research is essential for better innovation policies.

6.3 Optimal Academic Fraction of GDP

In this section we will look for the optimal public funding level for academic research as a fraction of GDP (R/Z) keeping the baseline subsidies fixed. This is particularly important because the rate of academic innovation is a major factor in determining the share of hot product lines, which determines the effectiveness of applied innovation.

The results indicate that welfare can be improved by allocating a larger fraction of GDP to

Table 11: Optimal Academic Funding (in percentages)

ψ	R/Z	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
10	0.8	22.0	0.54	0.76	85.4	0.50	0.80	4.5	8.9	10.0	1.37	95.4

academic research. In particular, when we consider only this as the policy tool, the optimal funding rate is 0.8% of GDP. This figure is only 0.5% in France (and in the benchmark case). Such a policy increases the fraction of hot product lines from 6.9% to 10%. However, this policy makes a limited contribution to growth and welfare due to the Ivory Tower nature of academic research.

So far we have assumed that public innovations have no immediate effect on productivity in a particular product line. However, one can argue that policies such as the Bayh-Dole Act, which was adopted in the US in 1980, enhance the applicability of academic innovations by allowing universities to retain ownership of inventions made using federal funds. This is an interesting policy question, which we can analyze in our setting. We will study this appropriability problem by considering a scenario where academic research is focused on immediately applicable innovations half of the time (Panel A), and one where all innovations are immediately applicable (Panel B). In the latter extreme, academic research functions much like private corporate research. It should be noted that academic research, as we have all experienced, has a much wider set of objectives than purely generating consumer products (such as education, to say the least). Our analysis will abstract from those considerations.

Table 12 summarizes the results of the optimal academic policy.

Table 12: Academic Funding with Alternative Bayh-Dole Scenarios (in percentages)

Panel A: Bayh-Dole=50%												
ψ	R/Z	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
10	1.9	21.7	1.35	0.94	84.7	0.38	1.9	4.4	8.7	16.4	1.49	96.3
Panel B: Bayh-Dole=100%												
ψ	R/Z	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
10	3.7	20.8	3.4	0.01	83.7	0.24	3.7	4.1	8.2	22.7	1.69	98.1

Under these alternative cases, the optimal level of academic funding is increasing in the applicability of academic research. While the optimal fraction is 1.9% when half of the innovations are immediately applicable, this number rises to 3.7% when all innovations are immediately applicable.

These optimal allocations bring with them large welfare gains, between 2 and 3 percentage points. Some of this is simply due to the increase in the Bayh-Dole factor, while the rest can be attributed to the optimal allocation of academic funding. The growth rate rises as well, attaining levels seen in the social planner's optimum in the last case.

Our results highlight the special role of academic research in overall growth and show the complementarities present between public and private research. Allocating resources to academic research has not only a direct effect on growth but also an indirect effect by making private research more productive. However, one should also note that this particular policy alone cannot make up for the underinvestment in research on the part of the private sector. Therefore, the next policy experiment is of particular importance.

6.4 Optimal Feasible Policy: Uniform Subsidy and Academic Budget

Our final policy experiment combines both of the feasible policies that have been considered thus far individually. We will allow both the uniform subsidy rate and the academic funding rate to be chosen by the policymaker. The advantage of considering both types of policies is to introduce more freedom to control the incentives for both types of research in a largely separate way. In particular, ψ and R/Z are going to be the choice variables in this exercise. The following table contains the results of this experiment

Table 13: Optimal Academic And Uniform Policy (in percentages)

ψ	R/Z	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
31	0.7	25.4	1.5	0.6	81.6	1.5	0.7	5.4	10.8	15.5	1.72	96.1

When considered jointly, the optimal uniform R&D subsidy is 31% and the optimal fraction of GDP allocated for public research is 0.7%. This combination generates a limited improvement, however. The growth rate increases 0.02 percentage points relative to the optimal uniform policy of Table 10 and 0.35 percentage point relative to the optimal public funding of Table 11. These improvements are mitigated by the limited applicability of the academic research. Next, we consider both uniform policy and academic funding jointly under the scenario where academic innovations have immediate applications for production.

Table 14: Academic And Uniform Policy with 100% Applicability (in percentages)

ψ	R/Z	τ_a	τ_b^e	τ_b^d	L_p	L_b	L_u	L_e	L_a	α	g	β
26	3.3	23.5	3.8	0.0	81.2	0.9	3.3	4.9	9.7	24.7	1.92	98.3

When academic innovations are geared toward consumer needs (i.e., have immediate application for production), the share of academic funding becomes much more effective. In this case, the optimal fraction of GDP allocated for academic research is 3.3%. Under this policy, the growth rate increases to 1.92% and achieves the highest welfare result among all policies considered. By using the level of academic funding to reach the proper share of researchers, the policymaker is able to lower the uniform subsidy, thus reducing needless cross-subsidization of applied research. Under the current policy 19% of the labor force is allocated to research, an increase over that of the baseline case. This time around, the composition of workers between applied and basic research is closer to the social optimum.

To summarize our findings, we first considered the most widely discussed policy, which is a uniform subsidy. Using this tool optimally yielded limited improvement in welfare due to oversubsidization of applied research since the policy could not distinguish between the research types with different spillover and productivity implications. Considering a policy combination that governs both private and academic research in which the researchers can appropriate the returns to their innovations could generate a significant improvement. The first main conclusion to be drawn for innovation policy is the importance of recognizing different types of innovations and the impact of policies on these types of research. The second is that it is important to take into account both the direct and indirect effects of academic research on productivity growth and the role of researchers' appropriability of their outcomes when considering growth and innovation policies.

7 Conclusion

In this paper, we distinguished between basic and applied research and identified spillovers associated with each. Our quantitative analysis highlighted the importance of this distinction. Indeed, in the competitive equilibrium, applied research is overinvested and basic research is underinvested. As a result, following a uniform research subsidy does not generate the expected welfare improvement due to inefficient cross-subsidization of applied research. An increase in the uniform subsidy improves the underinvestment in basic research by worsening the overinvestment in applied research.

The key message of our paper is that standard R&D policies can accentuate the dynamic misallocation in the economy. Our findings point to the need for policies that target basic research more directly. One method of achieving this is by increasing the intellectual property rights granted to academic researchers. Alternatively, one can reward collaboration between universities and the private sector, which would encourage focusing on research that can more directly lead to tangible gains in production technologies.

Our paper took a first step in trying to quantify the inefficiencies regarding different types of research and innovation efforts. There are still important open questions awaiting further study. In particular, the effect of university licensing and collaboration opportunities between universities and the private sector are some examples. We hope further structural work will be undertaken to enhance our understanding of the aforementioned issues, which can then guide the relevant innovation policies.

References

Acemoglu, D. (2008): *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2013): "Innovation, Reallocation and Growth," National Bureau of Economic Research WP 18993.

Acemoglu, D., U. Akcigit, D. Hanley, and W. R. Kerr (2012): "Transition to Clean Technology," University of Pennsylvania mimeo.

Aghion, P., U. Akcigit, and P. Howitt (2013): "What Do We Learn From Schumpeterian Growth Theory?," National Bureau of Economic Research WP 18824.

Aghion, P., and P. Howitt (1996): "Research and Development in the Growth Process," *Journal of Economic Growth*, 1(1), 49–73.

——— (1998): *Endogenous Growth Theory*. MIT Press.

——— (2009): *The Economics of Growth*. MIT Press.

Akcigit, U., and W. R. Kerr (2010): "Growth through Heterogeneous Innovations," National Bureau of Economic Research WP 16443.

Anselin, L., A. Varga, and Z. Acs (1997): "Local Geographic Spillovers between University Research and High Technology Innovations," *Journal of Urban Economics*, 42(3), 422–448.

Audretsch, D. B., and M. P. Feldman (1996): "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3), 630–640.

Bacchiocchi, E., and F. Montobbio (2009): "Knowledge Diffusion from University and Public Research. A Comparison between US, Japan and Europe Using Patent Citations," *Journal of Technology Transfer*, 34(2), 169–181.

Bikard, M. A. (2013): "Is Knowledge Trapped Inside the Ivory Tower? Technology Spawning and the Genesis of New Science-Based Inventions," MIT Sloan Mimeo.

Bloch, L., and E. Kremp (1999): "Ownership and Voting Power in France," FEEM WP 62.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81(4), 1347–1393.

Blundell, R., R. Griffith, and F. Windmeijer (2002): "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, 108(1), 113–131.

Dasgupta, P., and P. A. David (1994): "Toward a New Economics of Science," *Research Policy*, 23(5), 487–521.

Dhont-Peltrault, E., and E. Pfister (2011): "R&D Cooperation versus R&D Subcontracting: Empirical Evidence from French Survey Data," *Economics of Innovation and New Technology*, 20(4), 309–341.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001): "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, pp. 303–372. University of Chicago Press.

Garicano, L., C. LeLarge, and J. Van Reenen (2013): "Firm Size Distortions and the Productivity Distribution: Evidence from France," National Bureau of Economic Research WP 18841.

Goolsbee, A. (1998): "Does Government R&D Policy Mainly Benefit Scientists and Engineers?," *American Economic Review*, 88(2), 298–302.

Griliches, Z. (1986): "Productivity, R&D and Basic Research at the Firm Level in the 1970s," *American Economic Review*, 76(1), 141–154.

——— (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28(4), 1661–1707.

Henderson, R., A. B. Jaffe, and M. Trajtenberg (1998): "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988," *Review of Economics and Statistics*, 80(1), 119–127.

Howitt, P. (2000): *The Economics of Science and the Future of Universities.* Citeseer.

Impullitti, G. (2010): "International Competition and US R&D Subsidies: A Quantitative Welfare Analysis," *International Economic Review*, 51(4), 1127–1158.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108(3), 577–598.

Joint Economic Committee (JEC) (2010): "The 2010 Joint Economic Report," *Washington, DC: Government Printing Office*.

Klette, T. J., and S. S. Kortum (2004): "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 112(5).

Kortum, S. S. (1992): "Inventions, R&D and Industry Growth," Ph.D. Dissertation, Yale Univ., New Haven, CT.

——— (1993): "Equilibrium R&D and the Patent–R&D Ratio: US Evidence," *American Economic Review*, 83(2), 450–457.

Lentz, R., and D. T. Mortensen (2008): "An Empirical Model of Growth through Product Innovation," *Econometrica*, 76(6), 1317–1373.

Link, A. N. (1981): "Basic Research and Productivity Increase in Manufacturing: Additional Evidence," *American Economic Review*, 71(5), 1111–12.

Mairesse, J., and P. Mohnen (2010): "Using Innovations Surveys for Econometric Analysis," National Bureau of Economic Research WP 15857.

Mansfield, E. (1980): "Basic Research and Productivity Increase in Manufacturing," *American Economic Review*, 70(5), 863–873.

——— (1981): "Composition of R and D Expenditures: Relationship to Size of Firm, Concentration, and Innovative Output," *Review of Economics and Statistics*, 63(4), 610–615.

——— (1995): "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing," *Review of Economics and Statistics*, 77(1), 55–65.

Morales, M. F. (2004): "Research Policy and Endogenous Growth," *Spanish Economic Review*, 6(3), 179–209.

Nelder, J. A., and R. Mead (1965): "A Simplex Method for Function Minimization," *Computer Journal*, 7(4), 308–313.

Nelson, R. R. (1959): "The Economics of Invention: A Survey of the Literature," *Journal of Business*, 32(2), 101–127.

Pakes, A., and Z. Griliches (1984): "Estimating Distributed Lags in Short Panels with an Application to the Specification of Depreciation Patterns and Capital Stock Constructs," *Review of Economic Studies*, 51(2), 243–262.

Powell, M. J. (1970): "A Hybrid Method for Nonlinear Equations," *Numerical Methods for Nonlinear Algebraic Equations*, 7, 87–114.

Romer, P. M. (2001): "Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers?," in *Innovation Policy and the Economy, Volume 1*, pp. 221–252. MIT Press.

Rosenberg, N. (1990): "Why Do Firms Do Basic Research (with Their Own Money)?," *Research Policy*, 19(2), 165–174.

Rosenberg, N., and R. R. Nelson (1994): "American Universities and Technical Advance in Industry," *Research Policy*, 23(3), 323–348.

Segerstrom, P. S. (1998): "Endogenous Growth without Scale Effects," *American Economic Review*, 88(5), 1290–1310.

Trajtenberg, M., R. Henderson, and A. Jaffe (1992): "Ivory Tower versus Corporate Lab: An Empirical Study of Basic Research and Appropriability," National Bureau of Economic Research WP 4146.

——— (1997): "University versus Corporate Patents: A Window on the Basicness of Invention," *Economics of Innovation and New Technology*, 5(1), 19–50.

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(2), 431–436.

Zangwill, W. I., and C. Garcia (1981): *Pathways to Solutions, Fixed Points, and Equilibria.* Prentice-Hall Englewood Cliffs, New Jersey.

Appendix

A Theoretical Proofs

Proof of Proposition 1. Let $\mathcal{F}(\cdot, t)$ be the distribution over q at time t. Similarly, let $\mathcal{F}_H(\cdot, t)$ and $\mathcal{F}_C(\cdot, t)$ be the product type (hot/cold) conditional distributions. Thus, we will have $\mathcal{F}(q, t) = \alpha\mathcal{F}_H(q, t) + (1-\alpha)\mathcal{F}_C(q, t)$. The evolution of the aggregated productivity index $\bar{q}$ is then given by

$$\begin{aligned}
\bar{q}^{\varepsilon-1}(t+\Delta t) = &\int_0^\infty q^{\varepsilon-1} d\mathcal{F}(q, t+\Delta t) \\
=&\alpha \int_0^\infty q^{\varepsilon-1} d\mathcal{F}_H(q, t+\Delta t) + (1-\alpha)\int_0^\infty q^{\varepsilon-1} d\mathcal{F}_C(q, t+\Delta t) \\
=&\alpha \int_0^\infty \left[\Delta\tau\,(q+\eta\bar{q})^{\varepsilon-1} + (1-\Delta\tau)q^{\varepsilon-1}\right] d\mathcal{F}_H(q, t) \\
&+ (1-\alpha)\int_0^\infty \left[\Delta\tau_a\,(q+\lambda\bar{q})^{\varepsilon-1} + \Delta\tau_b^e\,(q+\eta\bar{q})^{\varepsilon-1} + (1-\Delta\tau)q^{\varepsilon-1}\right] d\mathcal{F}_C(q, t)
\end{aligned}$$

Thus the differential is

$$\begin{aligned}
\frac{\bar{q}^{\varepsilon-1}(t+\Delta t) - \bar{q}^{\varepsilon-1}(t)}{\Delta} =&\alpha \int_0^\infty \tau\left[(q+\eta\bar{q})^{\varepsilon-1} - q^{\varepsilon-1}\right] d\mathcal{F}_H(q, t) \\
&+ (1-\alpha)\int_0^\infty \left(\tau_a\left[(q+\lambda)^{\varepsilon-1} - q^{\varepsilon-1}\right] + \tau_b^e\left[(q+\eta)^{\varepsilon-1} - q^{\varepsilon-1}\right]\right) d\mathcal{F}_C(q, t)
\end{aligned}$$

and the normalized differential is

$$\begin{aligned}
\frac{\bar{q}^{\varepsilon-1}(t+\Delta t) - \bar{q}^{\varepsilon-1}(t)}{\Delta\bar{q}^{\varepsilon-1}(t)} =&\alpha \int_0^\infty \tau\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right] d\mathcal{F}_H(\hat{q}, t) \\
&+ (1-\alpha)\int_0^\infty \left(\tau_a\left[(\hat{q}+\lambda)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right] + \tau_b^e\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right]\right) d\mathcal{F}_C(\hat{q}, t)
\end{aligned}$$

Finally, the growth can be expressed compactly as

$$g = \frac{\alpha\tau\mathbb{E}_{\hat{q}}^H\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right] + (1-\alpha)\left(\tau_a\mathbb{E}_{\hat{q}}^C\left[(\hat{q}+\lambda)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right] + \tau_b^e\mathbb{E}_{\hat{q}}^C\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right]\right)}{\varepsilon-1}$$

This can also be rearranged into

$$g = \frac{\tau_a\left(\alpha\mathbb{E}_{\hat{q}}^H\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right] + (1-\alpha)\mathbb{E}_{\hat{q}}^C\left[(\hat{q}+\lambda)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right]\right) + \tau_b^e\mathbb{E}_{\hat{q}}\left[(\hat{q}+\eta)^{\varepsilon-1} - \hat{q}^{\varepsilon-1}\right]}{\varepsilon-1}$$

■

Proof of Proposition 2. The firm value, in general form, can be expressed as

$$\begin{array}{rl} & r\mathcal{V}_t(\hat{\mathbf{q}},m) - \dot{\mathcal{V}}_t(\hat{\mathbf{q}},m) \\ = & \max_{a,b} \left\{ \begin{array}{c} \sum_{\hat{q}\in\hat{\mathbf{q}}} \frac{1}{\varepsilon}\hat{q}^{\epsilon-1}\frac{Z_t}{M} - nw_t\left[h_a(a) + h_b(b) + \mathbf{1}_{(b>0)}\phi\right] \\ +na\left[\alpha\mathbb{E}^H_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\eta\right\},m\right) + (1-\alpha)\,\mathbb{E}^C_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\lambda\right\},m\right) - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \\ +nb\left(1+\rho_m\right)\left[\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\eta\right\},m\right) - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \\ +\sum_{\hat{q}\in\hat{\mathbf{q}}}\tau\left[\sum_{\hat{q}\in\hat{\mathbf{q}}}\left[\mathcal{V}_t\left(\hat{\mathbf{q}}\backslash\left\{\hat{q}\right\},m\right) - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right]\right] \\ +x\frac{m}{M}\left[\alpha\mathbb{E}^H_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\eta\right\},m\right) + (1-\alpha)\,\mathbb{E}^C_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\lambda\right\},m\right) - p'_m - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \\ +x\left(1-\frac{m}{M}\right)\left[\alpha\mathbb{E}^H_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\eta\right\},m+1\right) + (1-\alpha)\,\mathbb{E}^C_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}+\lambda\right\},m+1\right) - p_m - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \\ +n\kappa\left[\mathbb{E}_{\hat{q}}\mathcal{V}_t\left(\hat{\mathbf{q}}\cup\left\{\hat{q}\right\},m\right) - \mathcal{P} - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \\ +\kappa\left[n\mathcal{P} - \mathcal{V}_t\left(\hat{\mathbf{q}},m\right)\right] \end{array} \right\}. \end{array}$$

Now, conjecture $\mathcal{V}_t(\hat{\mathbf{q}}) = \frac{Z_t}{M}\left[\sum_{\hat{q}\in\hat{\mathbf{q}}} V(\hat{q}_t) + nV_m\right]$. When we substitute the conjecture into the the above expression and using the prices

$$\begin{array}{rcl} p_m & = & V_{m+1} + \mathbb{E}_{\hat{q},s}V\left(\hat{q}_{t+\Delta t} + \hat{s}\right) \\ p'_m & = & V_m + \mathbb{E}_{\hat{q},s}V\left(\hat{q}_{t+\Delta t} + \hat{s}\right) \end{array}$$

we find

$$(r-g)\,V_m = \max_{a,b}\left\{ \begin{array}{c} -\tilde{w}\left[h_a(a) + h_b(b) + \mathbf{1}_{(b>0)}\phi\right] \\ +a\left[\alpha V^H + (1-\alpha)\,V^C + V_m\right] \\ +b\left(1+\rho_m\right)\left[V^H + V_m\right] \\ +x\left(1-\frac{m}{M}\right)\left[V_{m+1} - V_m\right] \\ -\tau V_m + \kappa\mathbb{E}_{\hat{q}}V(\hat{q}_t) \end{array}\right\}.$$

and

$$V'(\hat{q}_t)g\hat{q} + \left[\tau + \kappa + r - g\right]V(\hat{q}_t) = \frac{1}{\varepsilon}\hat{q}^{\epsilon-1}.$$

Note that the last expression is a differential equation as a function of $\hat{q}$. Then

$$V(\hat{q}_t) = \frac{\hat{q}_t^{\epsilon-1}}{\varepsilon\left[r + \tau + \kappa + g\left(\epsilon - 2\right)\right]}.$$

This completes the proof. ■

Derivation of Multi-industry Distribution $\Gamma_{m,n}$.
We assume that when a firm loses its last product in a particular industry, it maintains a foothold there, in the sense that it still receives buy-out offers and can still directly use basic research relevant to that industry. When a firm loses all of its products or receives a destructive shock, it ceases to exist. We wish to find the joint distribution over the number of industries a firm is in and how many product lines it owns. For notational convenience, let us denote the basic research flow from m-industry firms by $\hat{b}_m = \mathcal{B}(\phi_m)b_m$. Let us also denote the expansion rate of a firm into a new industry by e_m. Here the expansion rate comes purely from buy-out offers by entrants. So given a per firm buy-out offer rate of x, a firm in m industries will expand at rate

$$e_m = x\left(\frac{M-m}{M}\right) = \left(\frac{\varsigma E a_e}{F}\right)\left(\frac{M-m}{M}\right)$$

Then the flow equation for firms in m industries with n products is

$$\begin{array}{rcl}
 & \text{Outflow} & \text{Inflow} \\
\left[\begin{array}{c} a_1+\hat{b}_1+\tau+\kappa \\ +e_1+\kappa \end{array}\right]\Gamma_{1,1} & = & a_e+2\tau\Gamma_{1,2} \\
\left[\begin{array}{c} a_m+\hat{b}_m+\tau+\kappa \\ +e_m+\kappa \end{array}\right]\Gamma_{m,1} & = & 2\tau\Gamma_{m,2}+e_{m-1}\Gamma_{m-1,1} \text{ for } m\geq 2 \\
\left[\begin{array}{c} n\left(a_m+\hat{b}_m+\tau+\kappa\right) \\ +e_m+\kappa \end{array}\right]\Gamma_{m,2} & = & \left\{\begin{array}{c} \left(a_m+\hat{b}_m\left(1-\rho_m\right)+\kappa\right)\Gamma_{m,n-1} \\ +3\tau\Gamma_{m,n+1}+e_{m-1}\Gamma_{m-1,n} \end{array}\right\} \text{ for } m\geq 1 \\
\left[\begin{array}{c} n\left(a_m+\hat{b}_m+\tau+\kappa\right) \\ +e_m+\kappa \end{array}\right]\Gamma_{m,n} & = & \left\{\begin{array}{c} (n-1)\left(a_m+\hat{b}_m\left(1-\rho_m\right)+\kappa\right)\Gamma_{m,n-1} \\ +(n-2)\rho_m\hat{b}_m\Gamma_{m,n-2} \\ +(n+1)\tau\Gamma_{m,n+1}+e_{m-1}\Gamma_{m-1,n} \end{array}\right\} \text{ for } n\geq 3,\ m\geq 1
\end{array}$$

where we use the convention $\Gamma_{m,-1}=\Gamma_{m,0}=0$ and $e_0=0$. The first line equates the outflows from $(m=1,n=1)$ that happen once the firm loses its product at the rate $\tau+\kappa$, acquires a new product line at the rate κ, innovates a new good at the rate $a_1+\hat{b}_1$ on average or expands into a new industry at the rate e_1. On the other hand, inflow happens from outsiders at the rate a_e and from the firms with 2 products that lose one of their products at the rate 2τ. Similar reasoning applies to the subsequent lines.

Using values for the $\Gamma_{m,n}$ distribution gives us the mass of firms in a given (m,n) state. The total mass of firms is then $F=\sum_{m=1}^{M}\sum_{n=1}^{\infty}\Gamma_{m,n}$. We ultimately want the mass of products in given industry state m. To get this we simply evaluate

$$\mu_m=\sum_{n=1}^{\infty}n\cdot\Gamma_{m,n}$$

■

B Data & Data Organization

Empirical investigation of the relationship between R&D investment and multi-market activity of a firm requires reliable and extensive information not only on product markets and on R&D characteristics of individual firms, but also on firm ownership status. The latter allows us to identify the product markets to which the firm is linked via its business group. We obtain this information from three different data sets.

R&D Information Information about R&D investment comes from the annual R&D Survey conducted by the French Ministry of Research. The R&D survey is available in annual waves of cross-sectional data, where the same firms are not necessarily sampled year after year (Mairesse and Mohnen (2010)). The survey covers a representative sample of French firms of more than 20 employees investing in R&D. However firms with less than .8 million euros of R&D investment fill out a shorter and simplified survey. The survey includes extensive information about the financing of R&D. It not only breaks down R&D investment according to the source of the funds but also provides its allocation to different types of R&D. More specifically, all firms are asked to report their R&D investment as either basic or applied research.

Multi-Market Activity The identification of business group structures is based on a yearly survey by INSEE called "Enquete Liaisons Financieres" (LIFI). It covers all economic activities but restricts its attention to firms that either employ more than 500 employees or generate more than 60 million euros in revenue, or hold more than 1.2 million euros of traded shares. However, since 1998 the survey is cross-referenced with information from Bureau Van Dijk and thus covers almost the whole economy. The LIFI survey contains information that makes it a unique data set for studying the relationship between multi-market activity and investment in basic research. Besides providing information on direct financial links between firms, it also accounts for indirect stakes and cross-ownership when identifying the head of the group. This is important as it allows us to precisely reconstruct the group structure even in the presence of pyramids. This feature allows us to obtain a reliable account of the structure of business groups in the French economy and, as a consequence, reliable measures of our key variable, the multi-market presence of business groups.

Since each firm can be active in several markets, we cross-reference the data set with an extensive yearly survey by the Ministry of Industry ("Enquete Annuelle des Entreprises"). The survey is filled out by French firms with more than 20 workers and contains information not only on the different markets in which a firm operates but also information on market dedicated sales for each segment. The data cover the vast majority of French firms and span the period 2000-2006.

Balance-Sheet Information We use the firm- and industry-level data sets based on accounting data extracted again from the EAE files. The data also include unique firm identifiers allowing us to match them to the R&D and LIFI data.

Data Organization

We first identify the ownership status of each firm in the economy and the head of the group with which the firm is affiliated. Indeed, our data source (LIFI) defines a group as a set of firms controlled, directly or indirectly, by the same entity (the head of the group). We rely on a formal definition of control, requiring that a firm holds directly or through cross-ownership at least 50% of the voting rights in another firm's general assembly. We do not expect this to be a major source of bias in our sample as most French firms are private and ownership concentration is strong even among listed firms.[17] Firms that do not conform to this definition are classified as stand-alone firms.

We then match the ownership information to our balance-sheet data and to our survey on lines of business within firms. We drop firms that appear in the ownership data but for which we cannot find balance-sheet information. We also delete as outliers firm-year observations whose ROA falls outside a multiple of five of the interquartile range and firms that report 0 employment or which have negative sales. Based on our two sources of information we identify the main line of business from the balance sheets and the different segments of the firm from the survey on lines of business. For computational convenience we create a new firm-group identifier that allows us to aggregate at the same time business groups, business groups with multi-divisional firms, exclusively multi-divisional firms and true stand-alone firms. We then define four measures of multi-market activity. The first measure counts each market in which the firm-group is present either via its ownership links or its multi-divisional structure. The second measure counts each market in which the firm-group is present with at least 9 employees via its ownership links or its multi-divisional structure. The third measure counts each market in which the firm-group is present exclusively via its ownership links. The final measure counts each market in which the firm-group is present exclusively via its ownership links and excluding financial activities.

We then define firm characteristics from balance-sheet data. There are three possible organizational types and comparison issues might arise. Taking the firm as the economic unit of interest has the advantage of simplicity since information is directly available in the balance sheets. However, this method has the disadvantage of not being comparable across organizational types. Indeed, most information for multi-divisional firms is aggregated across lines of segment, whereas firms belonging to business groups have market-specific information. Similar to existing studies by the Ministry of Research (Dhont-Peltrault and Pfister (2011)), we decided to aggregate the information to the economic unit at the highest level of control: the firm level for stand-alone and multi-divisional firms, and the business group level for firms affiliated through majority ownership.[18]

[17] In their overview of ownership structures and voting power in France, Bloch and Kremp (1999) show that ownership concentration is pervasive: for non-listed companies with more than 500 employees the main shareholder's ownership stake is 88%. The degree of ownership concentration is slightly lower for listed companies but still above 50% in most cases.

[18] In addition to the economic rationale for constructing the data at the highest level of control there is also a legal

In a final step we match the firms' balance-sheet and patent information to information contained in the R&D Survey. We focus on firms for which we have R&D information. Again we aggregate at the highest level of control. As before, one has to be cautious in aggregating on the basis of variables that might be prone to double-counting. When constructing information on the basic R&D intensity of a firm this is not the case as we are focusing exclusively on "internal" research expenditures. Therefore, if a member of the group contracts out research with another member of the group, then one will be counted as "external" research expenditures and the other one as "internal" expenditures. To correct for outliers in the dependent variable, we drop firm-year observations whose basic research intensity, conditional on positive basic research, falls outside a multiple of five of the interquartile range. In addition we exclude firm-year observations whose total R&D to sales ratio falls outside a multiple of five of the interquartile range.[19]

Variable List

All variables are organized and computed according to the method set out in the previous section. To summarize, we decided to aggregate the information to the economic unit at the highest level of control: the firm level for stand-alone and multi-divisional firms, and the business group level for firms affiliated through majority ownership. In the remainder of the document we will, for the sake of notational convenience, refer generically to firms.

- *Basic Research Intensity:* total basic research by firm i in year t divided by total applied research of firm i in year t. The formulation of the survey questions related to the type of research undertaken is directly derived from the definitions provided by the Frascati Manual;
- *# of Industries:* sum of all distinct SIC codes within firm i in year t irrespective of organizational form (business group or multi-divisional structure). Industries are successively defined at the 4-,3-,2- and 1-digit SIC levels;
- *# of Industries - Weighted Sum:* weighted sum of all distinct bilateral 1-digit SIC links within firm i in year t. Weights are computed on the basis of the empirical frequency of each bilateral SIC link in each year t;
- *# of Patent Classes Applied:* sum of cumulated distinct patent-class applications within firm i in year t. Cumulated patent-class applications are computed for the period leading from 1993 to year t. Patent classes are successively defined at the 5,4,3,2 and 1-digit levels (EPO Classification);
- *# of Patent Classes Granted:* sum of cumulated distinct patent-class grants within firm i in year t. Cumulated patent-class grants are computed for the period leading from 1993 to year t. Patent classes are successively defined at the 5-,4-,3-,2- and 1-digit levels (EPO Classification);
- *Financial Int.:* binary indicator equal to 1 if firm i in year t is present in a financial industry, 0 otherwise;
- *Foreign HQ:* binary indicator equal to 1 if the headquarters of firm i in year t are located outside France, 0 otherwise;
- *Market Share:* weighted average of total sales of firm i, year t in industry k divided by total industry sales in year t. Weights are computed on the basis of the industry share of employment within firm i in year t;
- *Outsourcing to Univ.:* binary indicator equal to 1 if firm i in year t has outsourced R&D to French universities, 0 otherwise;

argument. Indeed most public administrations and tribunals define the eligibility of firms for subsidy programs with respect to the business groups to which they belong.

[19] Alternatively, we exclude firm-year observations whose basic to applied R&D ratio falls above the 99^{th} percentile of the distribution. The results are qualitatively similar.

- *Profitability - ROA:* weighted average of EBIDTA divided by total fixed assets of all subsidiaries within firm i in year t. Weights are computed on the basis of the subsidiaries' share of employment within firm i in year t;
- *Profitability - ROS:* weighted average of EBIDTA divided by total sales of all subsidiaries within firm i in year t. Weights are computed on the basis of the subsidiaries' share of employment within firm i in year t;
- *Public R&D Funds:* binary indicator equal to 1 if firm i in year t has received French public funds, 0 otherwise;
- *Research Area:* weighted average of the share of respectively biotech / software / environmental research in research expenditures in firm i in year t. Weights are computed on the basis of the subsidiaries' share of total R&D within firm i in year t;
- *Total Employment:* total employment of firm i in year t;
- *IV - State Present in 1986:* binary indicator equal to 1 if the French state had a non-zero equity stake in firm i in 1986;
- *IV - SOE in 1986:* binary indicator equal to 1 if the French state had a controlling equity stake in firm i in 1986.

Online Appendix

OA-1 Theoretical Proofs

As the downstream production technology is unchanged in the generalized model and we continue to impose symmetry across the industries. This implies that

$$P_i = P = \frac{1}{M} \qquad \text{and} \qquad Y_i = Y = Z. \tag{OA-1}$$

Henceforth, we can drop the industry index i. The perfectly competitive firm that produces midstream good Y_i takes equilibrium prices P and p_j as given while maximizing its profit

$$\max_{y_j} \left\{ P \left[\int_0^1 y_j^{\frac{\varepsilon-1}{\varepsilon}} dj \right]^{\frac{\varepsilon}{\varepsilon-1}} - \int_0^1 p_j y_j dj \right\}.$$

This maximization leads to the following inverse demand for upstream good j

$$p_j = P \left(\frac{Y}{y_j} \right)^{\frac{1}{\varepsilon}}.$$

Monopolist in product line j, j has productivity q_j. The firm takes the demand function for its product as given and solves the following maximization problem

$$\pi_j = \max_{y_j} \left\{ P Y^{\frac{1}{\varepsilon}} y_j^{\frac{\varepsilon-1}{\varepsilon}} - \frac{w}{q_j} y_j \right\}$$

This delivers the following optimal quantity

$$y_j = \left[\frac{1}{M} \left(\frac{\varepsilon-1}{\varepsilon} \right) \left(\frac{q_j}{w} \right) \right]^{\varepsilon} Z$$

Plugging this into the production function for midstream goods, we find a relationship between wage w and aggregated productivity $\bar{q} \equiv \left(\int q_j^{\varepsilon-1} dj \right)^{\frac{1}{\varepsilon-1}}$

$$w = \frac{1}{M} \left(\frac{\varepsilon-1}{\varepsilon} \right) \bar{q} \tag{OA-2}$$

With this, we can greatly simplify the expression of the firm's quantity and price choices as a function of its normalized productivity $\hat{q}_j = q_j/\bar{q}$

$$y_j = \hat{q}_j^{\varepsilon} Z \qquad \text{and} \qquad p_j = \frac{1}{M \hat{q}_j}$$

Denote variables normalized by Z/M with a "$\sim$". Then the normalized profit and labor are given by

$$\tilde{\pi}_j = \frac{\hat{q}_j^{\varepsilon-1}}{\varepsilon} \qquad \text{and} \qquad l_j = \frac{\hat{q}_j^{\varepsilon-1}}{\tilde{w}} \left(\frac{\varepsilon-1}{\varepsilon} \right). \tag{OA-3}$$

where $\tilde{w}$ is the normalized wage. Note that by construction $\int \hat{q}_j^{\varepsilon-1} dj = 1$. As a result, we integrate OA-3 over j to find profit share and production labor share as

$$\frac{M \int_0^1 \pi_j dj}{Z} = \frac{1}{\varepsilon} \qquad \text{and} \qquad \frac{w L_P}{Z} = \frac{\varepsilon-1}{\varepsilon}. \tag{OA-4}$$

Finally, we combine OA-2 and OA-4 to find the final output as a function of aggregate productivity $\bar{q}$ and total production labor L_P :

$$Z = \bar{q} L_P / M.$$

Proof of Lemma 1 Let $\mathcal{F}_H(\cdot,t)$ and $\mathcal{F}_C(\cdot,t)$ be the aggregate product cumulative measures by type (hot or hold) at time t. For a small time step Δ, hot distribution $\mathcal{F}_H(\cdot,t)$ will satisfy

$$\begin{aligned}\mathcal{F}_H(\hat{q},t+\Delta) =& \mathcal{F}_H(\hat{q}/(1+\Delta g),t) - \Delta\tau\left[\mathcal{F}_H(\hat{q}/(1+\Delta g),t) - \mathcal{F}_H(\hat{q}/(1+\Delta g)-\eta,t)\right] \\ &+ \Delta\tau_b^e\mathcal{F}_C(\hat{q}/(1+\Delta g)-\eta,t) - \Delta\zeta\mathcal{F}_H(\hat{q}/(1+\Delta g),t) + \Delta\tau_b^d\mathcal{F}_C(\hat{q}/(1+\Delta g),t)\end{aligned}$$

Similarly, the cold distribution $\mathcal{F}_C(\cdot,t)$ will satisfy

$$\begin{aligned}\mathcal{F}_C(\hat{q},t+\Delta) =& \mathcal{F}_C(\hat{q}/(1+\Delta g),t) - \Delta\tau_a\left[\mathcal{F}_C(\hat{q}/(1+\Delta g),t) - \mathcal{F}_C(\hat{q}/(1+\Delta g)-\lambda,t)\right] \\ &- \Delta\tau_b\mathcal{F}_C(\hat{q}/(1+\Delta g),t) + \Delta\zeta\mathcal{F}_H(\hat{q}/(1+\Delta g),t)\end{aligned}$$

Finally, for $i \in \{H, C\}$, calculating

$$\dot{\mathcal{F}}_i(\hat{q}) = \frac{\mathcal{F}_i(\hat{q},t+\Delta) - \mathcal{F}_i(\hat{q},t)}{\Delta}$$

and taking the limit as $\Delta \to 0$ yields the desired flow equations. Note that for this we use

$$\frac{\mathcal{F}_i(\hat{q}/(1+\Delta g),t) - \mathcal{F}_i(\hat{q},t)}{\Delta} = -g\hat{q}[\partial\mathcal{F}_i(\hat{q})/\partial\hat{q}]$$

OA-2 Robustness Checks on Reduced-Form Results

In this section we provide further robustness checks on the correlation between a firm's basic research incentives and its multi-industry presence. Our baseline specification is related to the number of distinct 1-digit SIC activities in which a firm operates but extends to finer SIC classifications. All results are presented in Table OA-1.

Confounding Factors Columns (1) and (2) check robustness of the results with respect to confounding factors. Column (1) estimates the model only allowing for year and organization fixed effects, whereas column (2) includes a set of potential confounding factors. Results in column (1) suggest that presence in an additional industry, not accounting for other variables such as size, is associated on average with a 1.4 percentage-point higher basic research intensity of firms. In column (2) the set of regressors includes controls for size, profitability and headquarter localization. The impact of multi-industry presence is slightly lower but remains statistically significant.[20] Estimates on the localization of headquarters are also statistically significant at the 5% level. Total employment and profitability on the other hand are not.

Measures of Multi-Industry Presence Columns (3) and (4) provide alternative measures for the multi-industry presence of firms. Column (3) defines multi-industry presence on the basis of a firm's technological spectrum. To do so we use EPO patent data for French applicants. We define as the number of technology classes in which a firm is present as the cumulative distinct patent classes granted to the firm between 1993 and t. The coefficient is very similar in magnitude and precision to the one obtained using distinct 1-digit SIC industries. Column (4) measures multi-industry presence as a weighted sum of all distinct bilateral 1-digit SIC links within firm i in year t considering only distinct legal entities linked by majority ownership. Weights are computed on the basis of the empirical frequency of each bilateral SIC link in each year t. Intuitively, if a given bilateral industry link is rare, then industries are more likely to be very different. Multi-industry presence is still positively related to basic research intensity, the different point estimate being linked to the different support of the weighted industry variable.

[20] Further checks on control variables included market shares, R&D subsidies, collaborations with universities, the presence of financial intermediaries, state in the capital of the firm, industry fixed effects and the use of a mean patent scaling method.

Causality and Instrumental Variables Columns (5) and (6) address the potential concern of reverse causality, i.e. basic research leading to a larger economic scope of firms. We exploit historical ownership structures that affected a firm's multi-industry presence as instrumental variables. The two instruments are defined as *State Ownership 1985-1987* and *State Owned between 1985-1987*.

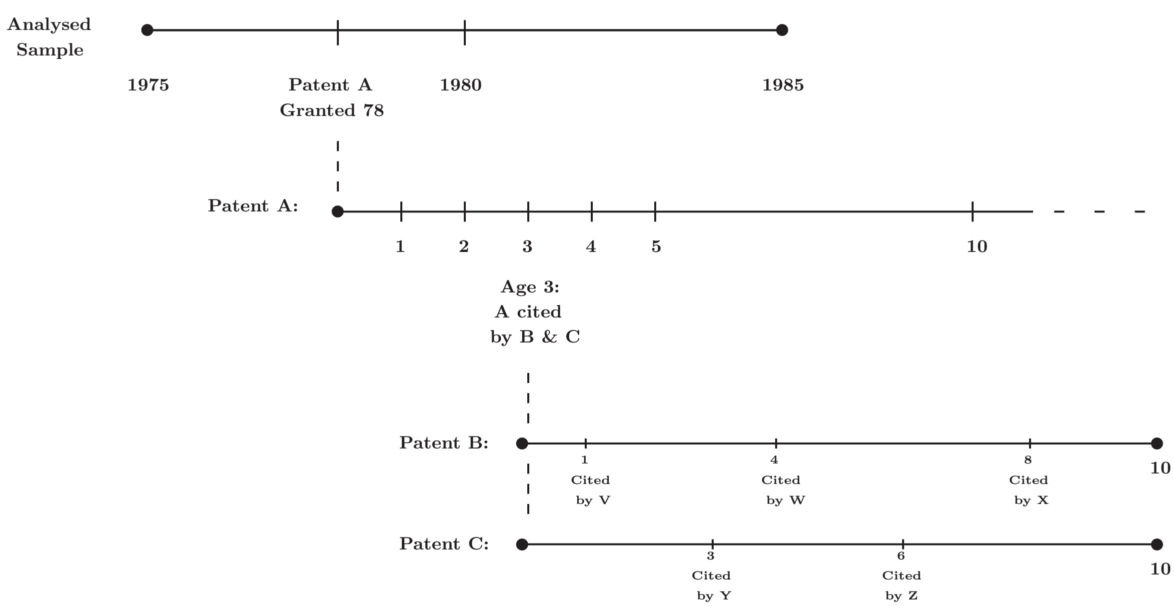
The rationale behind our identification strategy is as follows. In 1981 Francois Mitterrand was elected president of the Republic and implemented a vast nationalization program across industries. Even before that period the tradition of French state intervention resulted in a significant fraction of the economy being under state control. Consistent with Colbertist policies, the state also modified the economic scope of its firms by merging unrelated firms into large conglomerates of national champions. In 1987, however, Jacques Chirac was elected prime minister on a liberal platform and this marked the beginning of privatizations, which continued into the 90s. The embedded exclusion restriction therefore requires that state control in the 80 be associated today with a greater basic research intensity of firms only because of politically motivated mergers. The implicit assumption is that when these firms became private they adjusted their research spending from the social to the private optimum but did not adjust their multi-industry presence. First-stage estimates show that state ownership in the 80s is associated on average with 1.2 more industry links for firms between 2000 and 2006. The associated F-test are well above the critical levels related to weak instruments tables.[21] Columns 5 and 6 present the instrumented LATE coefficients related to multi-industry presence of the second stage. The coefficients are nearly twice as large in magnitude with respect to the non-instrumented coefficients of columns 1 and 2.

Estimation Columns (7) and (8) use alternative estimation methods for the baseline model with covariates. Column (7) presents estimates of the Heckman selection model, whereas column (8) presents estimates from a negative binomial model. In both cases estimates suggest a positive and statistically significant relation between basic research intensity and multi-industry presence.

OA-3 Construction of Within-Industry Spillovers

Figure OA-1 provides a graphical intuition for the computation of the citation information.

Figure OA-1: Computing the Cool-Down Rate



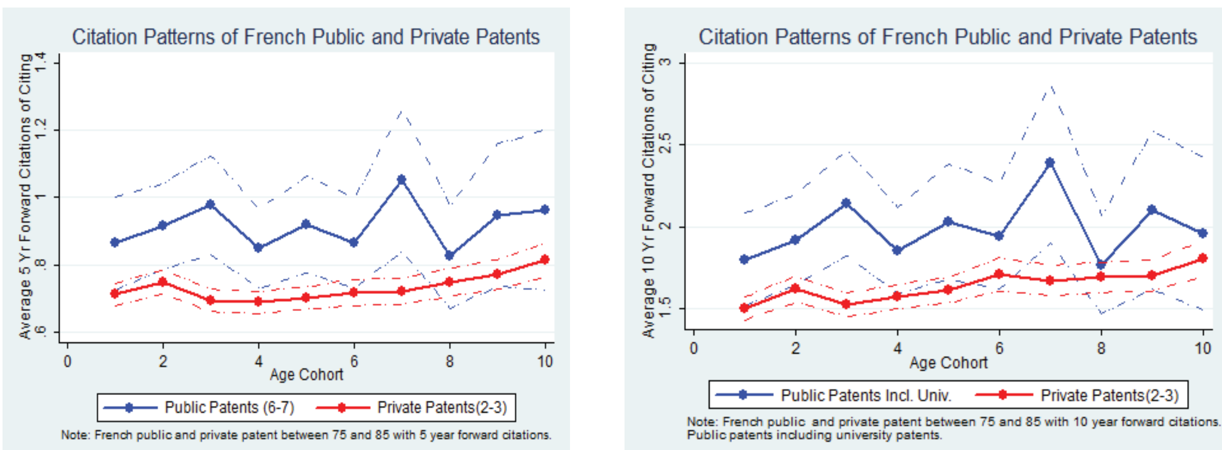
Patent A is granted in 1978. In 1981, when patent A is 3 years old, it receives citations from both patent B and patent C, which was applied for in 1981. Patent B in the following 10 years was

[21]The tables are available upon request.

cited by patents V, W and X, whereas patent C was only cited by patents Y and Z. The average citation of citing patents for patent A at age 3 is therefore 2.5. The timing of the computation implies that we need to be cautious with respect to possible truncation. We therefore compute our measure for patents between 1975 and 1985. This implies that, inclusive of the 10-years-forward lag, we can observe without truncation all our patents until the age of 10.

Robustness Checks Figure OA-2 provides robustness checks for the estimates on the cool-down rate of patents originating from basic and applied research. The left panel of the figure measures *Average Citations of Citing Patents* computing the 5-years-forward citations of the citing patents and is measured for patents granted in the period 1975-1985. The right panel re-classifies university patents that were defined as private depositors. In both cases results are unchanged, with a citation difference between public and private patents that becomes statistically non-significant at year 8. Indeed, in France, most of the academic patents are accounted for in the "public" category. French universities generally manage their patents through public research institutions with which academics are typically affiliated, one example being the CNRS.

Figure OA-2: Citation Patterns for French Public and Private Patents



Citation Differences for French Public and Private Patents

Age	1	2	3	4	5	6	7	8	9	10
5-Yr-Forward Citations	.15**	.16**	.28***	.16**	.22**	.15**	.33***	.08	.18	.15
	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)	(0.07)	(0.11)	(0.08)	(0.11)	(0.12)
10-Yr-Forward Citations	.3**	.3**	.62***	.28**	.42**	.23	.71***	.08	.39	.15
Including Univ.	(0.15)	(0.15)	(0.17)	(0.14)	(0.18)	(0.17)	(0.25)	(0.16)	(0.25)	(0.24)

*Note:*The figures separately plot *Average Citations of Citing Patents* for French public patents (blue line) and French private patents (red line) across patent age. The left panel computes *Average Citations of Citing Patents* computing the 5-years-forward citations of the citing patents and is measured for patents granted in the period 1975-1985. The bottom panel computes *Average Citations of Citing Patents* computing the 5-years-forward citations of the citing patents and re-classifying university patents as public patents. The table reports differences in citation patterns using two sample t-tests with unequal variances. One star denotes significance at the 10% level, two stars denote significance at the 5% level, and three stars denote significance at the 1% level

OA-4 Target Moments and Identification

In this section we explain the moments that are used to identify our parameters. For convenience, define expressions for the per product line R&D employment

$$h_a^m = \xi_b a_m^{\nu_b} \qquad \text{and} \qquad \bar{h}_b^m = \mathbb{E}_\phi \left[(\xi_b b_m^{\nu_b} + \phi) \cdot \mathbf{1}_{(\phi < \phi_m^*)} \right]$$

for applied and basic research, respectively. Note that these are functions of m, the number of industries a firm has working knowledge in.

Below, expectations are assumed to be over the distribution of firm characteristics $(m, n, \hat{\mathbf{q}})$. Note that here $\hat{m}$ denotes the number of industries in which a firm has one or more products, rather than the number of industries in which the firm has working knowledge. Since the latter is unobservable, we must compute the former to in order to match the data.

Basic Research Intensity by Number of Industries We define basic research intensity as the ratio of spending on basic research to spending on applied research. Since the effect of multi-industry presence on this quantity is of critical importance to our model, we have one moment for each $\hat{m} \in \{1, \ldots, M\}$. Given a set of parameters and an equilibrium of the model, this moment's value for a given $\hat{m}$ is

$$\Lambda(1-8) = \mathbb{E}_m \left[\frac{\bar{h}_b^m}{h_a^m} \middle| \hat{m} \right]$$

In our estimation, we use $M = 10$. However, in the data there are only a handful of firms with $\hat{m} > 8$, so we have one moment for each $\hat{m} \in \{1, \ldots, 7\}$ and a final moment which is averaged over $\hat{m} \in \{8, 9, 10\}$. The way in which this moment increases with $\hat{m}$ identifies the cross-industry spillover parameter p in our model. Additionally, it provides us with some identification power for the basic research cost parameters (ξ_b, ν_b).

Extensive Margin of Basic Research Investment by Number of Industries We use the share of positive basic research spending by each $\hat{m}$ to identify the mean μ and variance σ^2 of the fixed cost distribution basic research. This is simply the probability that the idiosyncratic fixed cost draw is less than the cutoff for a certain $\hat{m}$

$$\Lambda(9-16) = \mathbb{E}_{m,\phi} \left[\mathbf{1}_{(\phi < \phi_m^*)} \middle| \hat{m} \right].$$

Distribution of m We track two moments relating to the distribution of $\hat{m}$, the mean and mean squared. They are given by

$$\Lambda(17) = \mathbf{E}_{\hat{m}} \left[\hat{m} \right] \qquad \text{and} \qquad \Lambda(18) = \mathbf{E}_{\hat{m}} \left[\hat{m}^2 \right]$$

These moments identify the merger probability parameter governing the rate of expansion z as well as the mass of potential outside entrants (E). Together, these factors determine the equilibrium distribution of multi-industry presence.

Profitability Firm profitability is defined as the ratio of profits to sales. For a given panel of firms, this moment is given by

$$\Lambda(19) = \frac{1}{\varepsilon} - \mathbb{E}_{m,n,\hat{\mathbf{q}}} \left[\frac{\tilde{w} \left[h_a^m + \bar{h}_b^m \right]}{\frac{1}{n} \sum_i \hat{q}_i^{\varepsilon - 1}} \right]$$

Notice that there is one fixed component from static production side that yields information on the value of ε and another from dynamic R&D expenditures that yields information on R&D cost and step size parameters.

Exit Rate As exit occurs when firms either receive the exogenous destruction shock or lose their last product, the predicted exit rate will be

$$\Lambda(20) = \kappa + \tau \cdot \sum_{m} \Gamma_{m,1}$$

However, for consistency, we simply use the value from the simulated firm sample. This moment serves primarily to determine the value of the rate of exogenous destruction κ, as well as the mass of outside entrants E, since the size of the pool of entrants affects the rate creative destruction and hence the exit rate of single-product firms.

Total Research Intensity We have two moments to track levels of R&D: the ratio of total research labor expenditures spending to total production labor expenditures. Since research spending is proportional to n, R&D expenditures per product will be the same across firms with the same m, while employment will be a function of the portfolio of product qualities. Because the wage is common to both types of labor, this will simply be the ratio of R&D employment to production employment given by

$$\Lambda(21) = \mathbb{E}_{m,n,\hat{\mathbf{q}}} \left[\frac{\tilde{w} \left[h_a^m + \bar{h}_b^m \right]}{\left(\frac{\varepsilon-1}{\varepsilon}\right) \frac{1}{n} \sum_i \hat{q}_i^{\varepsilon-1}} \right]$$

Conditional on innovation rates, this moment give us information on the research production function parameters.

Firm Growth We have a moment for employment growth amongst firms. This is calculated conditional on the firm not exiting, since we do not observe the last period's growth rate for exiting firms. The moment is calculated by looking at the one-year growth rate of total employment by a firm. It is labeled $\Lambda(22)$. The employment growth primarily informs on the rate of exogenous destruction κ and the R&D cost function parameters.

Aggregate Growth The growth rate gives information on the effectiveness of research spending absent effects coming from the distribution of firm size and its relation to firm growth, particularly on innovation step sizes. This is moment $\Lambda(23)$.

Spillover Differential In order to quantify the spillovers associated with basic research, we turn to patent citation data. The model predicts that innovations that build off of previous basic research should have a larger step size on average. If we take citations as a proxy for step size, then patents that cite basic research should themselves have more citations.

This effect will diminish with the age of the patent due to product line cooldown. Thus the average time after which a public innovation is indistinguishable from a private innovation should be

$$\Lambda(24) = \frac{1}{\zeta} \left(\frac{\tau_a}{\tau} \right)$$

This yields direct information on the value of the cooldown rate ζ.

Firm Age Firm age is highly correlated with firm size. We track the average age of firms for those above and below the median firm size. This yields information entry and exit patterns, as well as on the rate of creative destruction. Moment $\Lambda(25)$ is the average age of firms below the median firm size, while moment $\Lambda(26)$ is the average age for those firms above it.

Table OA-1: Basic Research Intensity and Multi-Market Activity, Robustness Checks

	Covariates		Alternative Measures		Instrumental Variables		Estimation	
	(1) No	(2) Yes	(3) Patent Based	(4) Weighted Links	(5) State Present in 1986	(6) SOE in 1986	(7) Heckman	(8) Negative Binomial
# of Industries	0.014***	0.011***	0.012***	0.028***	0.023***	0.020**	0.045***	0.007***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Log Employment		0.002	0.003	0.018***	-0.003	-0.002	0.012***	-0.001
		(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Foreign HQ		-0.010**	-0.010	-0.048**	-0.004	-0.005	-0.041***	-0.006
		(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)
Profitability		0.006	0.016	0.005	0.005	0.005	0.025**	0.005
		(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
Year & Organization FE	YES	YES	YES	YES	YES	YES	YES	YES
N	13708	13706	3709	14823	13707	13707	13707	13707

Notes: Pooled data for the period 2000-2006. *Basic Research Intensity* is defined as the ratio of total firm investment in basic research divided by total firm investment in applied research. Columns 1 and 2 re-estimate the Tobit model with different sets of regressors. Columns 3 and 4 modify the measure of a firm's multi-industry presence. Column 3 uses patent applications of French firms to the European patent office (1993-2003) to count the number of distinct technological fields in which they are present (1-digit IPC classification). Column 4 weights each bilateral industry link of a firm by the empirical frequency of this link in the French economy, thus giving more weight to less related industries. Columns 5 and 6 re-estimate the model by instrumenting contemporary multi-industry presence by historical ownership structures. More specifically, we exploit the nationalization wave of the Mitterrand era that preceded the privatization of the 90s. The idea is that state ownership effectively increased the scope of a firm's economic activities. Column 5 uses state participation in the capital of a firm in 1986 as an instrument. Column 6 uses state ownership of a company in 1986 as an instrument. Both instruments accurately predict an increased multi-industry presence nowadays. Columns 7 and 8 estimate the relationship between multi-industry presence and basic research intensity by using a Heckman model and a negative binomial model. Tobit estimates relate to the marginal effect of the regressors with respect to the uncensored variable mean and are evaluated at the sample mean of covariates (except for categorical variables evaluated for firms that are present in 1 industry, non-foreign owned, in 2002). Robust standard errors clustered at the firm level in parentheses. See Appendix B for the definition of variables.