Dancing with the Stars: Innovation through Interactions*

Ufuk Akcigit Santiago Caicedo Ernest Miguelez

Stefanie Stantcheva Valerio Sterzi

March 27, 2025

Abstract

An inventor's own knowledge is a key input in the innovation process. This knowledge can be built by interacting with and learning from others. This paper uses a new large-scale panel dataset on inventors matched to their employers and patents from the European Patent Office. We document key empirical facts on inventors' productivity over the life cycle, inventors' research teams, and interactions with other inventors. Among others, most patents are the result of collaborative work. Interactions with better inventors are very strongly correlated with higher subsequent productivity. These facts motivate the main ingredients of our new innovation-led endogenous growth model, in which innovations are produced by heterogeneous research teams of inventors using inventor knowledge. The evolution of an inventor's knowledge is explained through the lens of a diffusion model in which inventors can learn in two ways: By interacting with others at an endogenously chosen rate; and from an external, age-dependent source that captures alternative learning channels, such as learning-by-doing. Thus, our knowledge diffusion model nests inside the innovation-based endogenous growth model. We estimate the model, which fits the data very closely, and use it to perform several policy exercises, such as quantifying the large importance of interactions for growth, studying the effects of reducing interaction costs (e.g., through IT or infrastructure), and comparing the learning and innovation processes of different countries.

JEL classification: O31, O33, O41, L16, H25

Keywords: Human Capital, Innovation, Interactions, Teams, Growth, Productivity, Inventors, Diffusion, Knowledge.

---

*Ufuk Akcigit, University of Chicago, NBER, and CEPR, uakcigit@uchicago.edu; Santiago Caicedo , Northeastern University, sa.caicedo@northeastern.edu; Ernest Miguelez, University of Barcelona, French National Centre for Scientific Research (CNRS), Bordeaux School of Economics (BSE), emiguelez@ub.edu; Stefanie Stantcheva, Harvard University, NBER and CEPR, sstantcheva@fas.harvard.edu; Valerio Sterzi, BSE, UMR CNRS 6060, University of Bordeaux, valerio.sterzi@u-bordeaux.fr. Jose M. Quintero, Marta Prato, Raphael Raux, Yulia Zhestkova, and Piotr Zoch provided excellent research assistance.

1 Introduction

Only a handful of ideas turn out to be transformative. Thomas Edison, Alexander Bell, or Nikola Tesla – all are familiar names thanks to the transformative inventions they introduced into our lives. When it comes to how these ideas are actually generated, the R&D investments behind them are only the tip of the iceberg. The inventor himself – with his own knowledge and human capital – is a key input. An entire process of knowledge accumulation by an inventor underlies the production of these new ideas. Inventors build their knowledge over time by interacting with others and learning from them. Universities, companies, and other research institutions try to create as many opportunities as possible for such interactions, from the informal common coffee room chat to formal international conferences. When it then comes to the production stage of ideas, inventors form research teams of varying sizes and use their knowledge to create ideas of heterogeneous qualities. These ideas lead to improved technologies and products, and, ultimately, translate into economic growth.

These observations lead to the following questions which we attempt to answer in this paper: What is the right theoretical framework that embeds interactions and knowledge diffusion at the inventor level into an innovation-based growth model? How can we micro-found the production of innovations at the level of inventors and research teams in a way that is disciplined by the data? How can we capture the processes through which inventors further develop their knowledge and productivity, which then translate into new ideas or innovations? Empirically, based on data on inventors, their interactions, and their innovations, what is the quantitative importance of interactions with others relative to other learning channels for individual productivity and growth?

To answer these questions we bring together the recent knowledge diffusion models and innovation-based growth models. Empirically, we need detailed data on who interacts with whom, on their preceding and resulting productivity, their innovation and research teams. While such data is very challenging to find, we make use of brand-new data that allows us to bring empirics to what is thus far a mostly theoretical literature on diffusion, namely panel data of the universe of inventors and patents registered in the European Patent Office. Therefore, our contributions are both theoretical and empirical.

In our theoretical contribution, presented in Section 2, we build a new model with learning and human capital formation through interactions (i.e., knowledge diffusion) nested inside an innovation-based growth model.[1] The existing literature has focused mostly on the role of either innovation or knowledge and human capital accumulation as explanations for economic growth. Our analysis highlights the importance of individual inventors' human capital for innovation, therefore combining these two important forces for economic growth into a single framework.

Viewed through the lens of the innovation-based growth literature, our model opens up the black box of ideas and micro-founds the innovation production process at the research team levels, where individual researchers' heterogeneous productivities evolve endogenously through interactions and knowledge diffusion. Viewed through the lens of the diffusion literature, the endogenous

[1] The knowledge diffusion model is inspired by Kortum (1997), Lucas (2009), Luttmer (2007), Lucas and Moll (2014), Perla and Tonetti (2014), and König, Lorenz, and Zilibotti (2016) among many others, while the innovation models are (Aghion and Howitt, 1992; Jones, 1995; Grossman and Helpman, 1991; Klette and Kortum, 2004; Aghion, Akcigit, and Howitt, 2014). We defer the detailed discussion of the related literature to Section 1.1.

interactions of researchers and knowledge diffusion contributes to their accumulated human capital (knowledge stock) or individual productivity. Knowledge itself is not directly used to produce the output, but instead is first used inside research teams to create an idea or innovation. These in turn improve aggregate productivity. Knowledge is diffused among others through interactions; and more knowledge leads to the production of higher-quality innovations.

Our empirical contribution is to bring new data to a mostly theoretical literature. We leverage new large-scale micro data that allows us to give empirical content to our theory. It is difficult and rare to find a dataset that tracks many individuals, their productivity, and their interactions over time. We manage to do so thanks to the new panel data of inventors and patents. Despite the fact that not all innovations are patented, our data carries information on millions of inventors who are responsible for a significant share of innovation around the world, which makes it very valuable for our purpose. A very recent special feature of our data is that inventors' names are uniquely disambiguated in order to turn the patent data into an inventor-level panel data. This major step allows us to observe and study the innovation and interaction histories of 826,878 inventors. In addition, inventors are carefully matched to their firms. There is also a compelling representation of numerous countries. In particular, we are able to quantify the productivity of individuals using the quality of their innovations over their life cycle. We can also measure interactions in several meaningful ways, such as past co-inventors, inventors in the same firm, and inventors in the same geographical region.

The data serves two major purposes in dialogue with our theory: First, it guides the key components of our theory and; second, it allows us to estimate the model and quantify the effects at play, and the impacts of changes. The model is disciplined by our main findings in the data – discussed in Section 3– which are used as building blocks:

1. Most patents are the result of collaborative work and produced by teams of heterogeneous sizes. The team size distribution is right-skewed.

2. Interactions among inventors are strongly correlated with higher subsequent inventor productivity (i.e., inventor innovation quality), even after including several detailed controls, such as inventor fixed effects. Inventors learn from each other to produce better innovations.

3. Interactions with inventors better than oneself are even more strongly correlated with higher subsequent productivity.

4. We observe a concave life-cycle profile of inventor productivity. This resembles the general earnings or wage life cycle profile documented by labor economists for other types of agents: First sharply increasing, then flattening. Through the lens of the model, inventors accumulate knowledge and become more productive over time, exactly according to this concave path.

5. Research productivity is positively associated with age, even after controlling for interactions with others and numerous controls. Thus, in our model, interactions should not be the only source of productivity improvements. We need to introduce an "external learning" channel, which serves as a catch-all for learning-by-doing, experience, or individual discovery.

Grounded in these empirical facts in our new data, our model can be viewed in more detail, but still at a single glance in Figure 1.

Building up from the individual inventor level (the bottom of the figure), inventors can learn, i.e., improve their productivity in two ways: They endogenously choose a meeting rate with others (interactions) at a search or effort cost, and they have access to an external, exogenous channel (which captures learning-by-doing, experience, individual discovery, etc.). In every period, a fraction of the inventors have new ideas to produce innovations, and form teams. To do so, they endogenously decide whether to become team leaders, who hire other inventors to their teams and reap the profits from innovation, or team members – earning the skilled researchers' wage. This endogenous occupational choice induces a realistic team structure that mirrors the data and which will match the empirical team size distribution. There are different learning opportunities for team members and team leaders.[2] Team members learn "within" teams by interacting with their team leader and more knowledgeable co-inventors, as well as from the exogenous learning channel. While, team leaders only learn outside their team from the exogenous source. Research teams produce innovations of heterogeneous qualities, increasing in the team leader's productivity and in the team size. These innovations are sold to final good producers, who use them to increase their TFP and produce output. The growth in aggregate productivity, or TFP, every period will be the total innovation quality produced by research teams.

In our quantitative contribution, we map the theoretical model back to the data and estimate it via the simulated method of moments and indirect inference based on 86 distinct moments from the micro inventor data. As described in Section 4, we obtain an extremely close fit to the data – despite being highly over-identified – and are able to match very well some non-targeted moments.

We then put our estimated model to use and perform four quantitative policy exercises in Section 5: Shutting down interactions vs. external learning, reducing the search costs of meeting others, and reducing access to outside information (excessive agglomeration). We find that interactions with others are quantitatively very important for improving inventors' productivity and, ultimately, for economic growth. In addition, there is a strong complementarity between access to external knowledge and learning from others: If others around oneself learn more from outside sources and then interact more, one will end up interacting with more knowledgeable people and learning more as well. We also estimate the model separately for the two major patenting countries: The U.S. and Germany. We show that these two countries have very different learning environments, which map into different lifecycle evolutions of research productivity, thus also confirming findings on the wage lifecycle profiles in these two countries (Lagakos, Moll, Porzio, Qian, and Schoellman, 2018b,a).

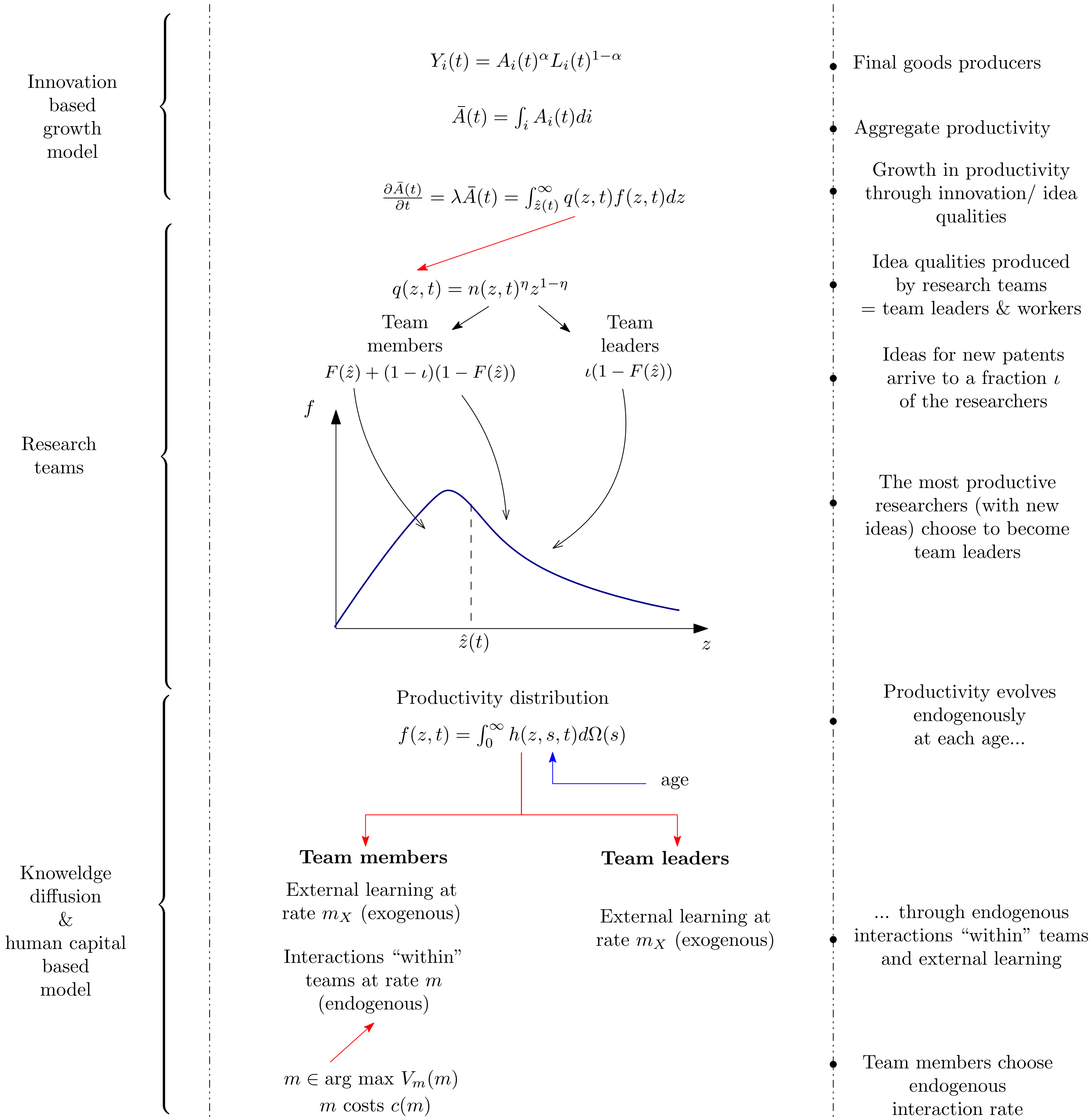
1.1 Related Literature

On the theoretical side, our paper combines two strands of the economic growth literature: Innovation-based models, and technology or knowledge diffusion models.

The innovation-based endogenous growth models of Aghion and Howitt (1992), Romer (1990), Jones (1995), Grossman and Helpman (1991), Klette and Kortum (2004) view innovation as the result

[2]This differential learning between team leaders and team members resembles the assignment of knowledge diffusion in Luttmer (2015).

Figure 1: Summary of the Model



of costly investments. Innovation improves aggregate TFP.[3]

In the technology or knowledge diffusion literature, the focus is mostly on imitation and technol-

[3] See Aghion, Akcigit, and Howitt (2014) for a survey of the creative destruction literature and Akcigit and Kerr (2018) for a recent example.

ogy adoption, rather than on the production of innovation per se. "Innovation" in these models is most often modeled as a draw from an exogenous distribution (or a stochastic shock to productivity) in contrast to "imitation" or "technology adoption," which are a draw from the endogenous technology distribution of firms (incumbents) or agents in the economy. In Jovanovic and Rob (1989), new ideas come from the interactions between individuals with different knowledge. Kortum (1997) builds a model of technological change where researchers sample ideas from a common distribution and pioneered the use of sources of ideas with fat tails that generate unbounded growth. Lucas (2009) introduces interactions between people who share ideas and a cohort structure. The starting distribution of productivities is unbounded and fat-tailed, so that even if agents only learn from each other, there can be sustained long-run growth.

More closely related to our work are the recent papers by Lucas and Moll (2014) and Perla and Tonetti (2014). Lucas and Moll (2014) allow for an endogenous resource allocation to production vs. interactions (i.e., imitation).[4] They characterize the extent to which search and imitation externalities reduce welfare. Staley (2011) extends the Lucas model by allowing each agent's productivity to experience stochastic shocks (termed "individual discovery"). Perla and Tonetti (2014) also features a trade-off for firms between producing and paying a search cost to upgrade one's technology through imitation, with the least productive firms choosing to imitate. Again, a thicker tailed productivity distribution leads to more growth (see also Benhabib, Perla, and Tonetti (2014)).

An interplay between innovation and technology diffusion is also explored also in Luttmer (2007, 2012, 2015). The second of these papers presents a growth model of imitation and selection where entrant and incumbent firms can imitate incumbents and there are Brownian shocks to productivity, which lead to firms remaining heterogeneous despite imitation.[5] In the third paper, agents can make stochastic individual discoveries, yet also learn from others. Agents with more knowledge are teachers, while the others are students; higher ability students are assigned to more productive teachers. The resulting assortative matching will lead to significant differences in earnings.

Most tightly related are the very recent papers by König, Lorenz, and Zilibotti (2016) and Benhabib, Perla, and Tonetti (2021). Benhabib, Perla, and Tonetti (2021) propose a model that does not use an infinite support assumption, but which features an endogenously expanding finite technological frontier through innovation. Firms choose between innovating, adopting technology, or producing. In König, Lorenz, and Zilibotti (2016), firms can improve their heterogeneous productivities through two channels: Expanding resources on producing innovations – which generate stochastic productivity improvements, or imitating other firms. Imitating firms are randomly matched to others and may succeed in adopting (i.e., imitating) their technology. Firms choose whether to innovate or imitate depending on their position in the productivity distribution, with firms further away from the frontier choosing to imitate. They obtain convergence to long-run growth even from limited initial heterogeneity: Firms who are close to the frontier choose to innovate, which allows them to draw a new productivity that does not come from the existing pool of ideas. In recent work, Caicedo,

[4]One of the focal points in the technology diffusion literature has been to what extent the models can endogenously generate unbounded growth. In these imitation models, since agents can only improve by imitating others around them if the learning opportunities are limited, growth will eventually stop. To avoid this, it is typically assumed that the initial productivity distributions have fat tails so that imitation (or interaction) opportunities never get exhausted.

[5]This is an alternative to the Lucas and Moll (2014) assumption that the stock of ideas is never exhausted.

Lucas Jr, and Rossi-Hansberg (2019) introduce learning in an economy with production hierarchies but cannot measure interactions or innovation directly. Also related to our work, Jarosch, Oberfield, and Rossi-Hansberg (2021) and Herkenhoff, Lise, Menzio, and Phillips (2024), study learning at the workplace and, consistent with our findings, infer strong coworker spillovers using wage data.

Diffusion and innovation across countries, as a function of their distance to the frontier, is also explored in Acemoglu, Aghion, and Zilibotti (2006) and Benhabib, Perla, and Tonetti (2014).

In the trade literature, Alvarez, Buera, and Lucas (2017) and Buera and Oberfield (2020) develop semi-endogenous growth models, in which trade favors the international diffusion of technologies (see also Alvarez, Buera, and Lucas (2008)). Perla, Tonetti, and Waugh (2021) study the effects of opening up to trade on growth in a model where technology diffuses, and heterogeneous firms can choose to adopt already existing technologies.

Our first main contribution is to bring empirical content to this class of models and is made possible by a new panel data on inventors. The data both guides the main features of our model and enables us to quantify it. Regarding the theory contribution, note that in these papers, it is directly the technology used in production that is diffused. Technology is improved either through innovation (modeled as external draws or productivity shocks) and/or imitation (modeled as draws from the endogenous distribution). Instead, in our model, the diffusion is of *knowledge* itself. Individual learning and discovery (external draws), and interactions (endogenous distribution draws) are two ways of diffusing knowledge. This is then used by research teams to produce ideas ("innovations"), which improve the technology of production and expand the finite technology frontier. Thus, the innovation-based growth models and diffusion models are truly linked and connected in our framework. We also model the micro-structure of the innovation production by introducing endogenously formed research teams as in the data, and a lifecycle of researchers.

On the empirical and quantitative side, we relate to several strands of the literature. The large and growing importance of teams for research has been documented in many empirical papers, recent examples of which are Wuchty, Jones, and Uzzi (2007), Jones, Wuchty, and Uzzi (2008), Azoulay, Fons-Rosen, and Zivin (2019), and Jaravel, Petkova, and Bell (2018). Jones (2009) in particular shows that "the burden of knowledge" over the past 50 years has increased and that, as a result specialization and teamwork have gained in importance. The relation between the life cycle stage (age) of inventors and research outcomes is also very important in the data, as documented by Jones (2010), who estimates shifts in life-cycle productivity of inventors and shows that they have become less productive at younger ages (see also Jones and Weinberg (2011)). We also are aligned with a copious empirical literature documenting technology spillovers (a prominent example of which is Bloom, Schankerman, and Van Reenen (2013)) by offering one possible explanation for why such spillovers may occur, namely, interactions between researchers. Our model can be used to study the effects of IT on the spread of knowledge and productivity as documented by Bloom, Sadun, and Reenen (2012) by modeling the latter as a reduction in the costs of interactions (see our analysis in Section 5). Finally, our setting can also provide a microfoundation (see Section 5) for the heterogeneous productivity and life cycle profiles observed among skilled workers across different countries as documented in Lagakos, Moll, Porzio, Qian, and Schoellman (2018b) and Lagakos, Moll, Porzio, Qian, and Schoellman (2018a).

2 Model

In this section, we develop a model of innovation and growth through learning and interactions. We first lay out the model and then solve for the equilibrium on a balanced growth path. Because it facilitates comprehension of the main forces and patterns, we at times solve for intermediate steps and optimizations as we go.

2.1 Brief Model Overview

Figure 1 gives a brief schematic overview of the model's structure. There are two sectors in the economy: A research sector that produces ideas (i.e., innovations) and a goods' production sector whose output is the final consumption good. There are two types of workers: skilled research workers who produce innovations and unskilled labor used in production, with no transition between these two types.

We start (at the top of the figure) from a standard innovation-based growth model in which aggregate quality $\bar{A}(t)$ evolves according to

$$\frac{\partial \bar{A}(t)}{\partial t} = \lambda \bar{A}(t)$$

where λ is the so-called "step size." This standard part is captured here by the goods production sector, in which final goods producers purchase the innovations (or ideas) from research teams and combine them with unskilled labor to produce the final good. Our contribution is to explicitly open up the black box of productivity growth and to model how the aggregate productivity improvement $\lambda \bar{A}(t)$ is produced through research teams in the research sector.

In the research sector, skilled workers of heterogeneous productivities decide whether to become team leaders or team workers; the most productive researchers become team leaders. Team leaders form research teams and hire team workers to produce an innovation or idea, the quality of which is a function of the team leader's productivity and the team size.

Finally, the heterogeneous productivities of research workers evolve endogenously over time as a result of learning, through two channels: Endogenous interactions, i.e., by meeting others and acquiring their knowledge, and an exogenous external learning source.

In this way, we combine the knowledge-diffusion (or human capital-based models) with the innovation-based growth models. Researchers acquire knowledge and diffuse it by meeting others, and that knowledge is used to produce innovations that push the technological frontier.[6]

[6] Thus, researchers' productivity and knowledge is different from the technology used in production of Lucas and Moll (2014) and the other papers described in Section 1.1. Here, productivity of researchers is research-related knowledge, which needs to go through the process of innovation by teams in order to be converted into technology, i.e., into an increase in aggregate productivity or TFP A.

2.2 Final Good Producers

There is a continuum of infinitely lived final good producers indexed by $i \in [0,1]$. Their production function is:

$$y_i(t) = A_i(t)^\alpha L_i(t)^{1-\alpha}$$

where L_i is production (unskilled) labor and A_i is their total factor productivity (TFP), A_i. The total final good produced for consumption in the economy is:

$$Y(t) = \int_0^1 y_i(t)di.$$

A final good producer can improve his factor productivity $A_i(t)$ by purchasing innovations of varying qualities from the research teams. The process through which research teams produce ideas is described in detail below. A final good producer who purchases an innovation of quality q from research teams can improve his current factor productivity $A_i(t)$ by the increment q. This "market for ideas" can be modeled in many different ways without affecting our results (see below). Without loss of generality, the price of the final good is normalized to 1. The wage of unskilled production workers is $w_u(t)$, and the total supply of unskilled workers is 1.

At any instant t, a final good producer with existing factor productivity $A_i(t)$ chooses a level of unskilled labor $L_i(t)$ by solving the (static) maximization problem $\max_{L_i}\{A_i(t)^\alpha L_i^{1-\alpha} - w_u(t)L_i\}$. This yields the optimal unskilled labor input choice:

$$L_i(t) = \left[\frac{(1-\alpha)}{w_u(t)}\right]^{\frac{1}{\alpha}} A_i(t).$$

The flow profits of the final good producer are thus:

$$\Pi_i(t) \quad = \quad \left[\frac{(1-\alpha)}{w_u(t)}\right]^{\frac{1-\alpha}{\alpha}} \alpha A_i(t).$$

From the first order conditions and using the labor market clearing for unskilled labor for all t, $\int_0^1 L_i(t)d_i = 1$, we also obtain that:

$$w_u(t) = \bar{A}(t)^\alpha(1-\alpha); \qquad L_i(t) = \frac{A_i(t)}{\bar{A}(t)} \qquad y_i(t) = \frac{A_i(t)}{\bar{A}(t)^{1-\alpha}}$$

where

$$\bar{A}(t) = \int_0^1 A_i(t)di$$

is total aggregate factor productivity (or, TFP). Aggregating across all producers, we obtain the total final good produced as a function of aggregate quality:[7]

$$Y(t) = \bar{A}(t)^\alpha.$$

[7] Note that if the level of unskilled labor were instead L, we would have $Y(t) = A(t)^\alpha L$.

The market for ideas on a balanced growth path

It will help to explain here how the market for ideas works, even though it requires anticipating some results presented later and assuming the economy is on a balanced growth path (which will be our focus throughout). The goal here is to show that it is irrelevant for aggregate productivity and growth in equilibrium, which final good producer purchases which team's innovation, and that the change in aggregate productivity will – under all surplus sharing rules – be equal to the sum of the qualities of the innovations produced by research teams.

We can show that, on a balanced growth in which aggregate productivity and the unskilled wage grow at constant rates, the value of any final good producer will be linear in his productivity $A_i(t)$. More precisely, in Appendix A-1, we demonstrate that the value of a final good producer with productivity $A_i(t)$, on a balanced growth path on which the wage $w_u(t)$ grows at a constant rate g_w and given interest rate r is:

$$V_i(t) = v\frac{A_i(t)}{\bar{A}(t)^{1-\alpha}} + \tilde{v}\bar{A}(t)^{\alpha} \qquad \text{where} \quad v \equiv \frac{\alpha}{\left[r + \frac{1-\alpha}{\alpha} g_w\right]}, \qquad \text{and} \quad \tilde{v} \quad \text{is constant.}$$

The surplus or change in value from buying an idea quality q from research teams is thus:

$$\Delta V_i(t) = \frac{v}{\bar{A}(t)^{1-\alpha}} q$$

This surplus has to be shared between the final good producer and the research team. In fact, the exact market structure in the market for ideas will be irrelevant. This is because the marginal value of increasing $A_i(t)$ is constant for all i, so that the return to an additional unit of productivity is the same across all final goods producers and independent of their current level of factor productivity $A_i(t)$. Thus, regardless of which final good producer purchases which team's innovation – and of how the surplus is shared – aggregate productivity growth will be the same.[8] We can for instance assume that final goods producers are randomly matched to research teams and that the research team receives a fraction β of the value generated. The price per unit of innovation quality is then simply: $p = \beta\frac{v}{\bar{A}(t)^{1-\alpha}}$. For all surplus sharing rules, the change in aggregate productivity equals the sum of innovation qualities produced by research teams. If research team j produces innovation quality q_j, then:

$$\frac{\partial \bar{A}(t)}{\partial t} = \int_i \frac{\partial A_i(t)}{\partial t} di = \int_j q_j dj = \int_{\mathcal{Z}^l} q(z,t) f(z,t) dz, \tag{1}$$

where $\mathcal{Z}^l$ denotes the set of team leader, $q(z,t)$ the quality of innovations produced by their team, and $f(z,t)$ the productivity distribution. In the next section, we provide a full description of these variables and explain how the quality of ideas q_j produced by research teams is connected to the productivity of team leaders.

This completes the link between the research sector – which develops, then sells innovations—

[8] Since there is a fixed number of researchers and no extensive margin switches between skilled and unskilled labor, the wage of skilled workers will scale with the price of innovation and the latter will not affect the optimal innovation intensity of research teams.

and the production sector – where those innovations are bought, then used to produce with higher factor productivity.

2.3 Learning and Research Teams

In the research sector, researchers first need an *idea* to produce innovations. We suppose new research ideas arrive to a fraction ι of individuals every instance. Once an idea arrives, individuals produce innovations of varying quality. At a given time t, all researchers have heterogeneous innovation productivities, $z(t) \in [0, \infty)$, distributed according to a cumulative density function $F(z, t)$.[9] Variable $z(t)$ can be called human capital or researcher knowledge. Researchers endogenously group into *research teams*, which consist of a *team leader* and n *team members* (or team workers). Each researcher inelastically provides one unit of labor. As the number of researchers is fixed, we normalize the mass of their total labor to 1. Only individuals with a new research idea can become team leaders. A team leader with productivity z who hires n members produces an idea of quality q, where:

$$q = z^{1-\eta} n^{\eta}. \tag{2}$$

The quality of ideas thus increases in the team leader's productivity and in the number of members. We use $\eta \in [0, 1]$ to denote the team leader's span of control (as in Lucas (1978)). When a research idea arrives to a team leader, teams are randomly formed out of the pool of potential team members and dissolved immediately after production. Let the wage of research team members at time t be denoted by $w(t)$. The price at which each unit of idea or innovation quality can be sold is $p(t)$.[10]

Inventors "learn" over time, i.e., improve their productivity z throughout their careers. This learning comes from two sources: through endogenous interaction with other individuals or through external, individual sources. The learning mechanism varies between team members and team leaders: team members endogenously learn from interactions with both team leaders and other more knowledgeable co-inventors, as well as from external sources, while team leaders only learn outside their research teams from the external or individual learning.[11]

External or individual learning: The external or individual learning channel captures the possibility of an individual improving his productivity even without meeting other individuals. At a general level, this is a catch-all term for all learning that occurs outside of interactions, such as learning-by-doing, learning through general work experience, reading, and individual discovery. This ingredient makes the model more realistic *per se*; in reality, individuals can learn on their own, and in the data, productivity is strongly correlated with age, even after controlling for interactions (see Section 4). In addition, introducing a learning channel other than interactions with others is key for the quantitative analysis. It allows us to empirically quantify the importance of each of the two learning channels for growth, without mechanically loading all productivity improvements on the interaction

[9] The productivity distribution will be endogenized below.

[10] This price is determined in the final goods market according to some surplus sharing rule, i.e., some market structure, as described in Section 2.2.

[11] This is akin to the measurement of interactions in the data (see section 3), where interactions are the cumulative number of more knowledgable coinventors within a research team.

channel. Put differently, we give the model the chance to show that interactions do not matter for productivity improvements.

Learning from the external source occurs through a Poisson process. With arrival rate m_X, an agent gets a productivity draw from an exogenous distribution of $E(z,s,t)$. For generality, the latter is allowed to depend on an agent's age s, as well as on calendar time t; how much an agent is able to learn from external sources can change over time and can be different for younger and older agents.

Endogenous learning through interactions: The second way in which individuals learn is by interacting with others. In equilibrium, team leaders are the most productive individuals in the team. This implies only team members have interaction opportunities 'within' their teams with either their team leader or other more knowledgeable team members. They learn at a Poisson arrival rate m from other individuals. Since teams are randomly formed and dissolve instantly after production, it is as if they draw from the full productivity distribution. When team member i meets individual j, he comes out of the meeting with the max productivity among them, i.e, after i meets j in the time interval Δt:[12]

$$z_i(t+\Delta t) = \max\{z_i(t), z_j(t)\}.$$

Team leaders choose the number of team members n that maximizes profits. Dynamically, they learn outside their research teams from external sources of knowledge. We can write the optimization problem for team leaders in a recursive form. The Hamilton-Jacobi-Bellman (HJB) equation for a team leader with productivity z and age s, in time period t is,

$$\rho V_l(z,s,t) = \max_{n\geq 0} \{p(t) z^{1-\eta} n^{\eta} - w(t) n\} + \underbrace{m_X \int_z^{\infty} \left(\iota V_l(\tilde{z},s,t) + (1-\iota) V_m(\tilde{z},s,t) - V_l(z,s,t)\right) e(\tilde{z},s,t) d\tilde{z}}_{\text{Gains from external learning}}$$
$$+ \underbrace{(1-\iota)\left(V_m(z,s,t) - V_l(z,s,t)\right)}_{\text{No new research idea}} + \delta(0 - V_l(z,s,t)) + \frac{\partial V_l(z,s,t)}{\partial t}, \tag{3}$$

where δ denotes an exogenous death rate, which implies that the age distribution is stationary, with cumulative probability distribution $\Omega(s) = 1 - e^{-\delta s}$ over $[0, \infty)$. A fraction ι of team leaders get a new research idea and stay as team leaders, otherwise they transition to become team members.

Note the solution to the profit maximization problem is static. The optimal number of members that a team leader with productivity z hires at time t, $n(z,t)$, the idea quality produced $q(z,t)$, and profits $\pi(z,t)$ are given by:

$$n(z,t) = \left(\frac{p(t)\eta}{w(t)}\right)^{\frac{1}{1-\eta}} z \; ; \qquad q(z,t) = \left(\frac{p(t)\eta}{w(t)}\right)^{\frac{\eta}{1-\eta}} z \; ; \qquad \pi(z,t) = p(t)\left(\frac{p(t)\eta}{w(t)}\right)^{\frac{\eta}{1-\eta}} (1-\eta) z. \tag{4}$$

More productive team leaders have more researchers in their team, produce better quality ideas, and make larger profits.

Team members learn through endogenous interactions within their research teams and also from external sources of knowledge. Meeting others is costly and requires time and effort. As team

[12] As in Lucas and Moll (2014) and the other papers described in Section 1.1, these meetings are not assumed to be symmetric, although this is a feasible extension.

members, individuals choose the rate at which they learn from others, m, optimally weighting the cost and benefits of interactions. The cost of achieving a meeting rate m is given by $C(m,s,t)$ and can depend on calendar time and age. They earn $w(t)$ while they are team members, only increasing their income when they become team leaders. Agents discount at a rate is ρ. We can write the HJB equation for a team member of age s in period t as,

$$\rho V_m(z,s,t) = w(t) + \max_m \left\{ \underbrace{m \int_z^\infty \left(\iota V_l(\tilde{z},s,t) + (1-\iota)V_m(\tilde{z},s,t) - V_m(z,s,t)\right) f(\tilde{z},t) d\tilde{z}}_{\text{Gains from interactions with others}} - \underbrace{C(m,s,t)}_{\text{Costs of interactions}} \right.$$
$$+ \underbrace{m_X \int_z^\infty \left(\iota V_l(\tilde{z},s,t) + (1-\iota)V_m(\tilde{z},s,t) - V_m(z,s,t)\right) e(\tilde{z},s,t) d\tilde{z}}_{\text{Gains from external learning}}$$
$$\left. + \underbrace{\iota \left(V_l(z,s,t) - V_m(z,s,t)\right)}_{\text{New research ideas}} + \delta(0 - V_m(z,s,t)) + \frac{\partial V_m(z,s,t)}{\partial t} \right\}. \quad (5)$$

We use these HJB equations to describe the occupational choice between becoming a team member or a team leader.

Occupational choice

An individual with productivity z, age s, and a new innovation idea can decide whether to become a team member or a team leader. If he chooses to be a member, he receives wage $w(t)$ in exchange for his unit of labor. On the other hand, if he chooses to be a team leader, he receives profits $\pi(z,t)$ from building a research team and producing ideas. Both team leaders and team members face the same external learning opportunities. However, only team members learn within their teams. To decide whether to become a team member or a team leader, individuals compare the discounted value of their expected earnings considering these alternative learning opportunities that come with each occupation.

An individual with productivity z and age s chooses to be a team leader if his value function is larger than the value he would receive as a team member, i.e., if $V_l(z,s,t) = \max\{V_m(z,s,t), V_l(z,s,t)\}$. Since $V_l(z,s,t)$ is an increasing function of z, and $\pi(0,t) = 0 < w(t)$, there exists a cutoff $\hat{z}(s,t)$ such that any individual $z(s,t) > \hat{z}(s,t)$ chooses to be a team leader, and any individual $z(s,t) \leq \hat{z}(s,t)$ chooses to be a member. Note this cutoff depends on age s, as the learning opportunities and, hence, the value functions for team leaders and team members also vary with age. For a given wage $w(t)$, the indifference condition determines the value of these age-dependent cutoffs,

$$V_m(\hat{z}(s,t),s,t) = V_l(\hat{z}(s,t),s,t) \ \ \forall s. \quad (6)$$

We can use these indifference conditions together with the labor market clearing condition to simultaneously solve for the age-dependent productivity cutoffs $\hat{z}(s,t)$ and the team member wages $w(t)$, setting the labor supplied by team members to be equal to the labor demanded by team leaders:

$$\int_0^\infty \underbrace{F(\hat{z}(s,t),t)}_{\text{Prefer to be team member}} + \underbrace{(1-\iota)(1-F(\hat{z}(s,t),t))}_{\text{No new ideas}} d\Omega(s) = \iota \int_0^\infty \int_{\hat{z}(s)}^\infty \left(\frac{p(t)\eta}{w(t)}\right)^{\frac{1}{1-\eta}} zf(z,t)dzd\Omega(s). \tag{7}$$

Team members are composed of researchers with productivity below the cutoff $\hat{z}(s,t)$ and the fraction $(1-\iota)$ of individuals above the cutoff. Conversely, team leaders are the ι fraction of researchers in each period above the cutoff.

2.4 Learning Dynamics

We now describe the evolution of the individuals' productivity over time. Let $\tilde{H}(z,s,t_b)$ denote the cumulative distribution of productivity of an individual of age s, born in period t_b, and $\tilde{h}(z,s,t_b)$ its corresponding density function. When individuals are born, they draw from a non-degenerate initial distribution $\tilde{H}(z,0,t_b)$, $\forall z,t_b$. Below we study the learning dynamics by describing the Kolmogorov Forward Equations (KFE) for inventors below and above the productivity cutoffs.

Inventors below the cutoff: Individuals of age s at time t with productivity below the cutoff $\hat{z}(s,t)$ choose to be team members. Every instant, team members are randomly matched with team leaders to form research teams. They have two sources of learning: i) interacting within these teams at Poisson rate $m(z,s)$ (that can depend on age s and productivity z) or ii) learning from an external source distribution at a rate m_X. The KFE describes the inflows and outflows of individuals from a given cohort born in period t_b and some productivity level z as they age,

$$\frac{\partial \tilde{h}(z,s,t_b)}{\partial s} = \underbrace{-m(z,s)(1-F(z,t))\tilde{h}(z,s,t_b) + m(z,s)\tilde{H}(z,s,t_b)f(z,t)}_{\text{Learning from team leaders + co-inventors}}$$
$$\underbrace{-m_X(1-E(z,s,t))\tilde{h}(z,s,t_b) + m_X\tilde{H}(z,s,t_b)e(z,s,t)}_{\text{External learning}}, \quad \forall z < \hat{z}(s,t), \ \forall s. \tag{8}$$

Equation (8) describes the dynamics of density function for team members with productivity below the cutoffs. The outflows account for the team members that learn from their team leaders or their more knowledgable co-inventors at the rate $m(z,s)$, as well as the ones that learn from the external source at a rate m_X. Conversely, inflows come from team members learning from other co-inventors or from external learning to exactly the productivity level z.

Inventors above the cutoff: Every period, $(1-\iota)$ fraction of the inventors above the cutoff are team members, as only a proportion ι inventors have new innovation ideas to become team leaders. These team members increase their productivity by interacting with more knowledgeable team leaders and co-inventors. As for the case of inventors below the cutoffs, all inventors learn from the external knowledge source. The productivity of a cohort of inventors $z(t) > \hat{z}(s,t)$, of age s at time t, born in period t_b evolves as,

$$\begin{aligned}\frac{\partial \tilde{h}(z,s,t_b)}{\partial s} = & \underbrace{\tilde{H}(\hat{z},s,t_b)(m(z,s)f(z,t) + m_X e(z,s,t))}_{\text{Inflow of team members with } z<\hat{z}} \\ & \underbrace{+m(z,s)(1-\iota)\left(\tilde{H}(z,s,t_b) - \tilde{H}(\hat{z},s,t_b)\right) f(z,t)}_{\text{Inflow of team members with } z\geq\hat{z}} \underbrace{-m(z,s)(1-\iota)\left(1-F(z,t)\right)\tilde{h}(z,s,t_b)}_{\text{Outflow of team members with } z\geq\hat{z}} \\ & \underbrace{-m_X\left(1-E(z,s,t)\right)\tilde{h}(z,s,t_b) + m_X\left(\tilde{H}(z,s,t_b) - \tilde{H}(\hat{z}(s,t),s,t_b)\right)e(z,s,t)}_{\text{External learning}}, \quad \forall z \geq \hat{z}(s,t),\ \forall s.\end{aligned} \tag{9}$$

We can use the change of variable $t = t_b + s$ to map age and birth data into calendar time. Let $H(z,s,t) := \tilde{H}(z,s,t_b)$ denote the cumulative probability distribution of an individual of age s at calendar date t (with density $h(z,s,t) := \tilde{h}(z,s,t_b)$). The cross-sectional age-conditional and unconditional productivity distributions are related through:

$$F(z,t) = \int_0^\infty H(z,s,t)d\Omega(s), \quad f(z,t) = \int_0^\infty h(z,s,t)d\Omega(s). \tag{10}$$

2.5 Balanced Growth Path

We now characterize the balanced growth path of the economy, on which the economy's growth rate is constant and all quantiles of the productivity distributions grow at the same rate g (following a "traveling wave" pattern).[13] Let de-trended productivity be denoted by $x = ze^{-gt}$. On the BGP, the age-conditional cumulative probability distribution is $\Gamma(x,s)$ (with density $\zeta(x,s)$) and the cross-sectional distribution is $\Phi(x)$ (with density $\phi(x)$) such that, for all $z,s,t \geq 0$:

$$\Gamma(x,s) = H(z,s,t) \qquad \text{and} \qquad \Phi(x) = F(z,t).$$

Assumption 1. *On the BGP, the exogenous learning distribution $E(z,s,t)$ is invariant, i.e., there is a distribution $\Psi(x,s)$ such that, for all t, s, and z, $\Psi(x,s) = E(z,s,t)$.*

The de-trended cutoff $\hat{x}(s) := \hat{z}(s,t)e^{-gt}$ is constant for each age s along the BGP, since all quantiles of the productivity distributions grow at the same rate g.

On a BGP the value functions of team leaders and team members, the real wages, and the real profits grow at the same rate $g(1-\eta)$,

$$v_l(x,s) = V_l(z,s,t)e^{-g(1-\eta)t}, \quad v_m(x,s) = V_m(z,s,t)e^{-g(1-\eta)t}, \quad \frac{w(t)}{p(t)} = \frac{w_0}{p_0}e^{g(1-\eta)t}, \quad \frac{\pi(t)}{p(t)} = \pi(x)e^{g(1-\eta)t},$$

where $\pi(x) := \left(\frac{p_0\eta}{w_0}\right)^{\frac{\eta}{1-\eta}}(1-\eta)x$.

[13] As has been discussed before (Lucas and Moll, 2014), there are no general existence or uniqueness theorems for partial differential equations systems, and our analysis does not focus on that, nor on computing the solutions to all possible initial distributions. Instead, we focus on the analysis of a balanced growth path.

Assumption 2. *On the BGP, the real cost of meetings is proportional to aggregate productivity $A(t)$ (which grows at rate $g(1-\eta)$). Thus, on the BGP, the nominal cost can be written as: $C(m,s,t) = c(m,s)e^{g(1-\eta)t}$.*

We further assume that $c(m,s) := \frac{\kappa(s)}{2}m^2$. The cost of meetings is expressed in units of the final good (the numeraire) like all other prices. For our quantitative analysis we estimate a linear approximation of $\kappa(s) = \max\{\kappa_0 + \kappa_1 s,\ \kappa_{min}\}$, with $\kappa_{min} > 0$ so costs have positive values only.[14]

On a BGP, team members choose a meeting rate $m^*(x,s)$ that equates the marginal costs and marginal benefits of interactions taking as given the growth rate of the economy and the productivity distribution[15],

$$m^\star(x,s) = \frac{1}{\kappa(s)} \int_{\max\{x,\hat{x}(s)\}}^{\infty} \left(\iota\, v_l(\tilde{x},s) - (1-\iota)v_m(\tilde{x},s) - v_m(x,s)\right)\phi(\tilde{x})\mathrm{d}\tilde{x}\ , \quad \forall s. \tag{11}$$

Team members will choose a higher meeting rate if the value and probability of becoming a team leader is larger. For team members below the cutoff, $x \leq \hat{x}(s)$, the optimal meeting rate only depends on their age, while for team members above the cutoff, the benefits from interacting with others are smaller as they acquire more knowledge. In the limit, as the productivity of team members goes to infinity, they have no incentives to interact, $\lim_{x\to\infty} m^*(x,s) = 0$, $\forall s$.

We see from equation (4) and the fact that individual productivity grows at rate g that idea quality grows at rate $(1-\eta)g$. It follows immediately that the growth rate of aggregate productivity is $\frac{\partial \bar{A}(t)}{\partial t}\frac{1}{\bar{A}(t)} = g(1-\eta)$ and the growth rate of aggregate output is thus $\frac{\partial Y(t)}{\partial t}\frac{1}{Y(t)} = \alpha g(1-\eta)$.

Let $Z(t)$ denote average productivity at time t. Letting i index skilled researchers, we can write the proportional change in average productivity in a small time interval Δt, $\frac{Z(t+\Delta t)-Z(t)}{Z(t)\Delta t}$ and obtain the growth rate of productivity on the BGP by letting Δt go to zero, i.e.:

$$\begin{aligned} g &= \lim_{\Delta t\to 0} \frac{Z\,(t+\Delta t) - Z\,(t)}{Z\,(t)\,\Delta t} \\ &= \int_0^1 \left[\begin{array}{c} m(z_i,s_i) \times [F(\hat{z}_i,t) + (1-\iota)(1-F(\hat{z}_i,t))] \times \left[\frac{1}{Z(t)}\left[z_i F\,(z_i,t) + [1-F\,(z_i,t)]\,\mathbb{E}_F\left(z_i'\,(t) \mid z_i'\,(t) > z_i\right)\right] - 1\right] \\ +m_X \times \left[\frac{1}{Z(t)}\left[z_i F\,(z_i,t) + [1-F\,(z_i,t)]\,\mathbb{E}_E\left(z_i'\,(t) \mid z_i'\,(t) > z_i\right)\right] - 1\right] \\ +\delta \times \left[\frac{1}{Z(t)}\mathbb{E}_0(z_i'(t)) - 1\right] \end{array} \right] di \end{aligned} \tag{12}$$

where $\mathbb{E}_F$, $\mathbb{E}_E$, and $\mathbb{E}_0$ denote the expectations relative to the, respectively, cross-sectional, external, and age zero distributions.

We now summarize these characteristics of the BGP in the following definition.

Definition 1. ***Balanced Growth Path.*** *A balanced growth path (BGP) consists of a constant g, a set of invariant distributions $(\Phi(x), \Gamma(x,s))$, a set of value functions $(v_l(x,s), v_m(x,s))$, paths for the real wage $\frac{w(t)}{p(t)}$, team sizes $n(z,t)$, idea qualities $q(z,t)$, profits $\pi(z,t)$, cutoff for productivities $\hat{x}(s)$, and interaction rates $m(x,s)$, such that:*

(i) All quantiles of the invariant distributions grow at the same rate g, i.e.:

[14] The analysis does not depend on these parametric specifications, but we use this particular one to illustrate the results and to estimate the model later on.

[15] See the detailed derivation of the BGP equations in the Appendix A-1.2.

$\Gamma(x,s) = H(z,s,t) \ ,\forall z,t,s \geq 0$
$\Phi(x) = F(z,t) \ ,\forall z,t,s \geq 0$, *with density* $\phi(x)$.

(ii) $\Gamma(x,s)$ *and* $\Phi(x)$ *satisfy the KFEs* (8), (9) *and* (10) *for all* (x,s).

(iii) The growth rate satisfies (12).

(iv) The real wage $\frac{w(t)}{p(t)}$ *that clears the labor market for researchers satisfies* (7)

(v) Profits $\pi(z,t)$ *(as well as team size* $n(z,t)$ *and innovation qualities* $q(z,t)$*) are given by* (4).

(vi) The cutoff $\hat{x}(s)$ *is a solution to* (6).

(vii) The meeting rate $m(x,s)$ *solves* (11).

(viii) Aggregate productivity $\bar{A}(t)$ *grows at rate* $(1-\eta)g$.
Aggregate output is $Y(t) = \bar{A}(t)^{\alpha}$ *and grows at rate* $\alpha g(1-\eta)$.

We can, in addition, derive a useful relation between the economy's growth rate, the rate of endogenous interactions $m^*(s)$ (for team members $x \leq \hat{x}(s)$), and the external learning rate m_X. To do so, we assume that, on the BGP, the external and the age zero productivity distributions have a common tail parameter, θ:

Assumption 3. *Assume that* $\Psi(x,s)$ *and* $\Gamma(x,0)$ *have a common Pareto tail, i.e.*

$$\lim_{x\to\infty} \frac{1-\Psi(x,s)}{x^{-1/\theta}} = \varrho(s) \ , \ \lim_{x\to\infty} \frac{1-\Gamma(x,0)}{x^{-1/\theta}} = k_0$$

for some $\theta, k_0 > 0$, *and* $\varrho(s) > 0 \ \forall s$.

Note that the external source distribution has a location parameter $\varrho(s)$ that depends on age.

Suppose we start from a BGP with an initial distribution $F(z,0) = \Phi(x)$, where $\lim_{z\to\infty} \frac{1-F(z,0)}{z^{-1/\theta}} = \lim_{x\to\infty} \frac{1-\Phi(x)}{x^{-1/\theta}} = k < \infty$. We can show the cross-sectional productivity distribution $\Phi(x)$ has a Pareto tail θ (so $k > 0$) and derive an expression for the growth rate g,

Proposition 1. *Along a BGP, the cross-sectional productivity distribution* $\Phi(x)$ *has a Pareto tail, i.e.:*

$$\lim_{x\to\infty} \frac{1-\Phi(x)}{x^{-1/\theta}} = k > 0$$

and the following relation holds:

$$1 = \int_0^\infty e^{-\frac{g}{\theta}s} \int_0^s m^*(\tau)\Gamma(\hat{x}(\tau),\tau) e^{\frac{g}{\theta}\tau} d\tau d\Omega(s) + m_X \int_0^\infty e^{-\frac{g}{\theta}s} \int_0^s \frac{\varrho(\tau)}{k} e^{\frac{g}{\theta}\tau} d\tau d\Omega(s) + \frac{k_0}{k} \int_0^\infty e^{-\frac{g}{\theta}s} d\Omega(s) \tag{13}$$

Proof. See Appendix A-1. □

All else equal, this relation shows that the growth rate is increasing in the average rate of interactions of team members (below the cutoffs), the mass of team members $\Gamma(\hat{x}(s),s)$, the rate of external learning m_X, the scale of the productivity distribution at birth relative to the cross-sectional distribution, k_0/k, and the thickness of the productivity distributions' tails as captured by θ. It is decreasing in the death hazard rate δ. Naturally, in general equilibrium, as individuals endogenously respond by changing their occupation, $\hat{x}(s)$, and the intensity of interactions, $m^*(s)$, the effect of the model parameters on growth can also change. We further discuss these endogenous effects in our quantitative exercises below.

Let us consider a few key special cases. To compare the growth relation (13) to the previous diffusion literature, suppose there is no team structure, and all individuals produce autonomously.[16] In the model, this happens when idea production only depends on individual productivity and all inventors interact with each other at a fixed exogenous rate $m^*(x,s) = m, \ \forall x, s$. If the location parameter of the external distribution does not depend on age, i.e., $\varrho(s) = \varrho$ for all s, the growth rate implied by equation (13) is:

$$g = \left[(m + \frac{\varrho}{k} m_X) - \delta \left(1 - \frac{k_0}{k} \right) \right] \theta.$$

In addition, if the scale parameters of the external distribution and the age zero distributions are the same, $\varrho(s) = k_0 = k$ for all s, the growth rate depends only on the sum of the rates of interaction and external learning, and on the tail parameter:

$$g = (m + m_X)\theta.$$

Moreover, if the external distribution Ψ were not growing, the growth rate would simply be determined by the tail parameter and the endogenous interaction rate:[17,18]

$$g = m\theta.$$

When is there positive growth? A major focal point in the diffusion literature is whether sustained long-run growth can occur. It is important to note that our model is rich enough to allow for many different cases, with positive or zero growth. This is why we start from the most general case and cover several special cases above. Ultimately, the model will be disciplined empirically in Section 4. For the sake of completeness, we draw out all the possible cases that can arise with respect to the growth rate here:

(i) If either one of the initial productivity distribution $F(z, 0) = \Phi(z)$ or the external distribution $\Psi(z, s)$ has a fat tail, i.e., if $\theta > 0$, there is positive growth.[19]

(ii) If neither of these distributions has a fat tail ($\theta = 0$), and the external distribution is not growing, then there is no sustained growth in the long-run.

---

[16] For a recent review of the diffusion literature, see Buera and Lucas (2018).

[17] Although our model does not nest and is not nested in Lucas and Moll (2014), their growth rate most closely corresponds to our special case with $k_0 = k$ and no external ideas source or an external idea source that grows at rate $g_e < g$ strictly smaller than g. In this special case of our model, $g = m_I\theta$, as in Lucas and Moll (2014). This same growth rate also arises in an extension in Lucas and Moll (2014), in which there is a fixed (not growing) source of external learning. The rate of arrival of external ideas m_X is irrelevant, justifying the authors' claim that it is really interactions that drive growth. This is also true in our model as shown here, in the case where the external distribution is not growing or is growing at rate $g_e < g$. We allow for $g_e = g$ here, to ensure a fair quantitative comparison between the external learning source and the endogenous interactions. If we did not do this, we would always conclude mechanically that in the long-run, the external distribution loses its importance.

[18] Note also that, as in Lucas and Moll (2014), if the external distribution has a fat tail with tail parameter $\zeta > \theta$, then the growth rate would be $g = m\zeta$ (even if the external distribution were not growing at rate g).

[19] This is true even if the external distribution is not growing.

(iii) If the external distribution is growing at a rate g_e, growth can be sustained regardless of the shape of the distributions Φ and Ψ.

3 Data, Measurement and Empirical Findings

In this section, we give empirical content to the theory presented in Section 2. One of our main contributions is to leverage a new large-scale micro dataset – the European Patent Office data – that allows us to give concrete empirical counterparts to the concepts of innovations, research teams, interactions, and productivity of the model.

We first describe the data and empirically document some key facts about researchers and their teams, which were already used to inform the key elements of the model in Section 2. These facts also will be used as the main data moments to estimate the model in Section 4. To not repeat the figures twice, we will point the reader directly to each relevant figure highlighting a significant data moment in the estimation Section 4. One of the core facts of our model, which we study in this section, is the link between interactions and productivity improvements.

3.1 The European Patent Office Data

There are many patent offices in the world. Typically, academic papers on productivity have been based on patents filed with the United States Patent and Trademark Office (USPTO) (made available first by Hall et al. (2001)). Instead, our data is derived from the European Patent Office (EPO) data – an under-explored and novel source.

Patent data is ideal for studying the effect of interactions on learning and innovation. First, patent documents contain a multitude of information on the patent assignee (the original owner of the patent, which can be a firm, an individual, or a university), all the inventors who contributed to the innovation, and the innovation itself. Second, it allows us to track individual inventors over time. To do so, the first step is to disambiguate inventor names, in order to turn the list of inventors into an inventor level panel data. Disambiguation consists of identifying two or more inventors listed on several patents as the same person. This is based on their homonymy or quasi-homonymy (identity or similarity of names and surnames), together with related information (e.g., residential address, co-inventors, citations, or technological classes). We describe the disambiguation algorithm in Appendix OA-1. Because the EPO data was only very recently disambiguated, we have improved upon the disambiguation over time.[20]

A fundamental reason to use the EPO data is that it provides a better representation of different countries. This mitigates the "home bias" in patenting. Indeed, a patent only protects an innovation within the geographical jurisdiction of the patent office where it is filed. The USPTO would thus be preferred by U.S. firms (which can be seen in the large representation of U.S. firms' patents in the USPTO data). A patent in the EPO can come from any country in Europe, as well as from other patenting countries interested in exporting to the European market. Figure A-2 panel (A)

[20] The raw data used is the CRIOS-PatStat dataset described in Coffano and Tarasconi (2014) and Pezzoni et al. (2014). See also Breschi et al. (2016) and Akcigit et al. (2016).

depicts the number of EPO applications per country for the major patenting countries. The U.S. is the largest contributor to innovations at the EPO, followed by Germany, Japan, other European countries (France, Great Britain, Italy, the Netherlands, and Switzerland), and South Korea.

Overall, our data contains 2,955,055 applications. There are fewer patent applications at the EPO than at the USPTO (see Table OA-1 for a summary of the number of worldwide applications, as well as the number of applications to the EPO and the USPTO per year). Panel (B) of Figure A-2 also shows that the number of EPO patents has steadily increased over the years.

The disambiguated EPO data contains 3,474,514 unique inventors, of which 1,145,185 (32.96%) are listed in two or more applications.[21] The average number of patents per inventor is 2.2. However, patent production, like other scientific or academic endeavors, is a very skewed phenomenon. A substantial proportion of inventors has only one patent and very few super-productive inventors have many patents. Figure 2 shows the distribution of inventors with no more than 20 patent applications (99.33% of the sample). The majority (around 91% of the total) produce 1 to 4 patents during their career. Only 80,000 (2% of the sample) of inventors have more than 10 patent applications.[22]

We assign patent applications to specific years according to their "priority year," i.e., the first year of application to any patent office worldwide (thus, the priority date may differ from the application date at the EPO). This procedure ensures that the patent is assigned to the year closest to the actual innovation. We assign patents to origin countries using the inventors' country of residence.[23] Finally, each patent is also assigned a technology field. We adopt the OST (2008) classification of IPC codes into 30 fields. Online Appendix OA-1 provides more information on IPC codes and their technology classification, as well as summary statistics on the distribution of patents across fields in the EPO data.

3.2 Measuring Productivity, Interactions, and Team Composition

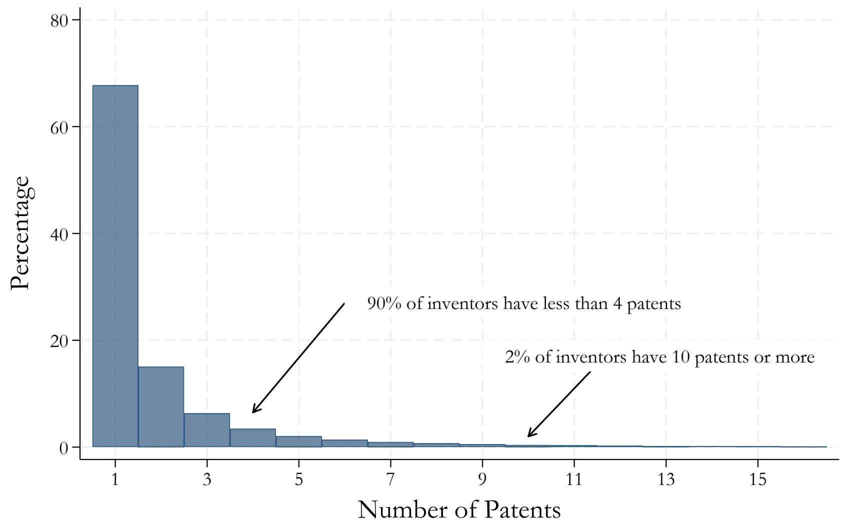
Our empirical analysis sheds light on the key channel highlighted by the model, namely on how interactions with others contribute to human capital accumulation and enhanced productivity. We now map each of the model's concepts of teams, team leaders, individual productivity, idea quality and interactions to the data. In doing so, we remain as true as possible to the spirit of the model. It is worth noting outright that the data is very rich and that there are several meaningful ways of measuring these variables. In each case, we provide several possible robustness checks using alternative definitions.

[21] This is very comparable to the USPTO disambiguation by Li et al. (2014), which identifies 3,124,041 unique inventors, out of which 37.39% (1,168,208) have filled more the one patent.

[22] Figure OA-1 shows that there is some variation in the number of patents per inventor across the top 10 patenting countries, ranging from 1.5 per inventor in South Korea to 3.1 in Germany. These differences reflect the differential propensities to work in teams, as well as the propensity to file with the EPO.

[23] There are different methods for assigning patents to countries. Traditionally, the applicant's address reported in the patent document has been used (usually, the firm's address). However, this approach can give a biased measure of where the actual innovation was developed, since many firms may assign their patents to their headquarters rather than to the branch where the innovation was produced. A more meaningful alternative approach – which we use– is to use the addresses of the inventors listed in the application. When inventors reside simultaneously in several countries – a relatively rare case– we can assign the patent probabilistically to each country.

Figure 2: Number of Patents per Inventor at the European Patent Office



Notes: The figure shows the percent of inventors with a given number of patents in their lifetime. Sample contains 3,474,514 unique inventors from patents filed with the EPO between 1977 and 2010.

Inventor teams

As in the model, research teams play a key role in the data. The patent applications at the EPO reflect the increasing importance of teams in patent production. In the time around 2010, more than 70% of patents were produced by teams with at least two inventors. Figure 3 depicts the distribution of team size and highlights that teams of one, two and three inventors account for the majority of patents. The average team size is 2.6 inventors; conditioning on multi-inventor patents, the average team size is 3.4 inventors.

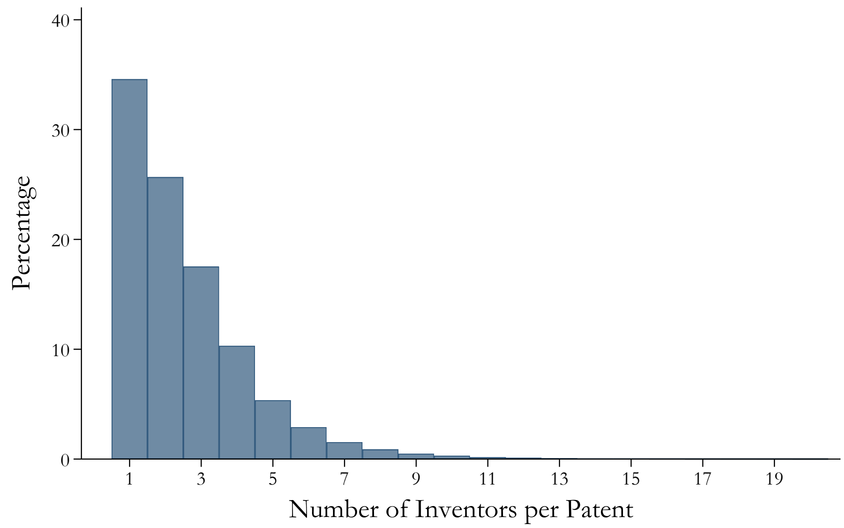
Idea quality

To measure the quality of the idea produced by a team, we use the quality of patents produced by the team. As is standard in the literature (see Hall et al. (2001); Pakes (1986); Schankerman and Pakes (1986); Trajtenberg (1990)), patent quality is measured by the number of forward citations. We count the citations a given patent receives in a time window of three years after the priority date. We use the three year window to account for truncation in the citations, i.e., the fact that more recent patents have had less time to accumulate citations, as described in Hall et al. (2001). Thus, our benchmark measure of the idea quality of team j at time t, $q_{j,t}$, is the citations of the patent(s) produced by team j at time t:[24]

$$q_{j,t} = \sum_{\tau=t}^{t+2} \text{citations}_\tau .$$

[24]If the same team produces two patents in the same year, we sum the citations to both of these patents in a 3 year window.

Figure 3: Team Size Distribution



Notes: The figure shows the team size distribution, i.e., the number of co-inventors per patent in the EPO data, based on 3,474,514 unique inventors from patents filed between 1977 and 2010.

An improvement we make over earlier work is that we consider forward citations coming from patents filed with the EPO, as well as all patents from a family in which one patent is filed with the EPO.[25] As a robustness check, we also provide results using a window of five years, as well as removing self-citations (see Table 3).

Panel (B) of Figure 5 shows the idea quality q of teams as a function of the team leader's age, i.e., the life cycle of idea qualities of team leaders. This life cycle profile will be one of the moments targeted in the estimation.

Individual productivity measures

To map the individual inventor's productivity in the model to its counterpart in the data at each time, we compute a set of different measures. These capture slightly different notions of productivity and allow us to check that our results are robust to various inventor productivity definitions. One natural measure of current (or flow) productivity of inventor i in year t, denoted by $p_{i,t}$, is the number of citations-weighted patents produced by inventor i at time t:

$$p_{i,t} = \sum_j q_{j,t}.$$

This accounts for both innovation quantity and quality produced by an inventor. The cumulative (or stock) productivity of inventor i at time t is his citations-weighted patent stock to date, $P_{i,t}$:

$$P_{i,t} = \sum_{s=t_0}^{t} p_{i,s}.$$

[25]On patent families, see **?**.

This stock measure takes into account the quantity and quality of an inventor's past innovations, in addition to current ones. This is our benchmark measure of an individual inventor's productivity.

Additional quality measures used in the analysis – some reported as robustness checks, others available upon demand – are the unweighted number of patents produced at time t (which captures only the quantity of innovation, regardless of quality), as well as average citations per patent (received within a 3-year window) of the patents produced at time t (which captures purely the quality of innovation, regardless of quantity). We also experiment with different truncation windows (e.g. five years) and with counting or excluding self-citations.

Panel (C) of Figure 5 shows the life cycle profile of productivity for all inventors, according to our benchmark measure. It is in line with those found in the recent literature on life cycle wages and productivity (Lemieux, 2006; Lagakos et al., 2018b): It is increasing and concave. We will use the productivities at different ages as moments to target in the estimation of the model in Section 4.

Assigning team leaders

Our theory suggests that the team leader is the most productive inventor. There are several ways of identifying in the data who is "the most productive" inventor of a team.

Our first and main method for picking the patent team leader is to simply compare productivities, as defined above, and to choose the team member among all inventors on a given patent who had the highest cumulative productivity in period $t-1$. Thus, inventor i will be the team leader of the team that files patent j if

$$P_{it-1} = \max_{n} \left(P_{n,t-1} \right),$$

where n indexes all inventors in the team.[26] The remaining inventors of the team, if there are any, are considered to be the "team members" equivalent of the model.

We also explored two alternative and complementary ways of identifying the team leader: As the most senior inventor on the team, or as the first inventor listed in the application. Our results (not shown) are robust to these alternative team leader assignments.

Measuring interactions

There are many possible ways to measure interactions – one of the key variables of the model. We show results for several different measures going from least to most broad, which are:

1. Past co-inventors.

2. Inventors in the same firm.

3. Inventors in the same geographical region.

The first measure only counts as interactions past co-inventors, i.e., inventors listed on the same patents at some point in the past. This is a strong definition of interactions since these are people one has worked with certainty. Also, it likely underestimates total interactions: Inventors have probably

[26]According to this measure, there could be more than one team leader for a given patent.

interacted with more people than just those with whom they have filed patents. To be exactly in line with the model, we could also count only past co-inventors who were better than the focal inventor. This measure of interactions will be our benchmark which corresponds to the model's idea that one learns only from better people, which we test and confirm below. The second and third definitions are broader. They count inventors who were better than oneself at the time of the encounter, and with whom one has been in the same firm (definition 2), or with whom one has lived in the same region (definition 3), at some point over one's past career. These are broader measures of interaction in the sense that one cannot be sure of the intensity of direct contact that actually happened with someone working in the same firm or living in the same region. The results below show, as is intuitive, that the effects on productivity from more interactions increase progressively as one moves from the weakest interaction measure (measure 3) to the strongest one (measure 1).

Regarding our benchmark measure, formally, we can count the interactions with co-inventors better than oneself, and worse than oneself, that one has had in the past. The high-quality interactions of inventor i at time t are defined as the number of unique co-inventors j of inventor i in all past years, who had a productivity $P_{j,s}$, larger than that of inventor i in the period right before the encounter or interaction took place:

$$\text{High Quality Interactions}_{i,t} = \sum_{s=t_0}^{t} \text{unique more productive co-inventors}_{i,s}.$$

Conversely, "Low Quality Interactions" are defined as the number of unique co-inventors of inventor i in all past years, who had a productivity $P_{j,s}$, lower than that of inventor i in the period right before their encounter. As a robustness check in Table 3, we also count repeat interactions, i.e., if inventor i interacts with inventor j in two different years, that interaction is counted twice.

The regressions also include a number of important additional controls, such as the team leader's age, which is the number of years he has already spent in the sample, the square of age, and controls for team characteristics, such as the team size, i.e., the number of inventors listed in a given patent. We also control for firm size using the number of inventors active in the firm at each time. To assign active inventors to their firms, we reconstruct the employment histories of all inventors based on the assignees (firms) listed on their patents.[27] The strategy used to build inventors' employment histories is described in Appendix OA-1.[28]

The sample consists of all inventors and patents between 1977 and 2010.[29] The first year of each inventor is omitted from the regressions in order to be able to construct the necessary quality and interaction measures. Table 1 presents summary statistics of the variables used in the regression analysis.

[27] For similar approaches, see Hoisl (2007) and Nakajima et al. (2010).

[28] The assignee, could also be a university, hospital, research centre, or even an individual inventor. We treat them all as the inventors' de facto employers for our purposes, which makes most sense for the variables we need to construct.

[29] We remove the patents after 2010 to minimize the truncation issue in citations.

Table 1: Summary Statistics

	Mean	Standard Deviation	Min	Max
Idea Quality				
Conditional on patenting:				
3-year citations	1.4	3.09	0	401
5-year citations	2.2	4.53	0	421
3-year citations (excluding self citations)	1.2	2.76	0	401
Unconditional on patenting (at individual level):				
3-year citations (all inventors)	0.3	2.30	0	589
3-year citations (team leaders only)	0.7	3.99	0	589
3-year citations (team leaders only: benchmark sample)	1.0	5.09	0	589
Interactions				
High Quality Interactions	1.7	3.25	0	68
Low Quality Interactions	8.9	18.0	0	578
Total interactions	10.6	19.9	0	605
High Quality Interactions in the firm	568.4	1538.5	0	31425
High Quality Interactions in the region	17295.8	57639	0	743616
Team and firm characteristics				
Team leader age	5.8	5.23	1	34
Team size	3.1	2.08	1	99
Firm Size	466.2	880.9	1	5930

Notes: Summary stats based on 1,560,713 inventor-year observations used in the regressions, corresponding exactly to the sample used in the regressions in Tables 2, Column 1. “Conditional on patenting” refers to the quality of patents produced by inventors. “Unconditional on patenting” refers to the annual productivity of inventors, measured as the weighted sum of the number of patents filed and the three-year citations those patents receive. In years when inventors do not file any patents, their productivity is considered to be zero. Statistics are computed for (i) all inventors, (ii) only team leaders, (iii) only team leaders in the benchmark sample. “Total interactions” is the sum of High and Low Quality interactions, as defined in the text. “Firm size” is the number of active inventors in a firm at a given time. See the text for a detailed description of these variables.

The effect of age

Intuitively, the effect of the external learning channel on productivity will be disciplined in the data by the effect of age. Importantly this effect must be measured while controlling for interactions and conditional on many detailed fixed effects. The age effect then captures the learning that happens over time, not through interactions, but through learning-by-doing, individual discovery and learning, or exposure to outside knowledge sources. The effect of age is consistently positive and significant. The lifecycle of productivity and idea quality is concave as was described before, and as can be seen in panels (B) and (C) of Figure 5.

3.3 The Link Between Interactions and Productivity

We now present our key results on the link between past interactions and subsequent productivity, controlling for characteristics of the team leader, the team's inventors, the firm, and the research environment.

Specification and estimation

Our baseline specification is as follows:

$$q_{j,t} = \beta_1 \text{Interactions}_{i[j],t-1} + \beta_2 Age_{i[j],t} + \beta_3 Age^2{}_{i[j],t} + \beta_4 \text{Team Size}_{j,t} + \beta_6 \text{Firm Size}_{j,t} \\ + B_t + B_{s[j]} + B_{f[j]} + B_{r[j]} + B_{r[j]} \times B_t + B_{s[j]} \times B_t + B_{s[j]} \times B_{r[j]} + B_{i[j]} + \varepsilon_{j,t}, \qquad (14)$$

where $q_{j,t}$ is the standardized (z-score) idea quality of team j at time t. The team leader of team j is denoted by $i[j]$, employed by firm $f[j]$, in region $r[j]$ and technology sector $s[j]$. Idea or innovation quality is measured in turn by one of the several measures described above and transformed into a z-score by subtracting the mean and dividing by the sample standard deviation to yield $q_{j,t}$. "Interactions" is a measure of the past interactions of the team leader; recall that we explore several measures. Age is the age of the team leader. With some abuse of notation for simplicity, B_t denotes a set of year-fixed effects, $B_{s[j]}$ are sector-fixed effects, $B_{f[j]}$ are firm-fixed effects, and $B_{r[j]}$ are region fixed effects.[30] In order to absorb more of the inventor and team heterogeneity, we add region-year, sector-year, sector-region, and, most importantly, individual inventor fixed effects.

Inventor fixed effects, in particular, go a substantial way towards alleviating a potential endogeneity problem, namely that more productive inventors may be able to both interact with more people and have more co-inventors. Conditional on inventor fixed effects, a positive coefficient on lagged interactions could still arise because of endogeneity, but in much less likely cases. It would have to be that an inventor who expects to be better than his own average productivity in year t intentionally starts interacting with more people at time $t-1$. Alternatively, there would have to be

[30] Regions r are defined at comparable administrative levels across countries. For instance, for the US, we use the BEA economic areas, which are equivalent to Metropolitan Statistical Areas (MSAs) in the case of metropolitan regions. For Europe, we use the Nomenclature of Territorial Units for Statistics (NUTS3) level. For all remaining countries, we use the smallest level used by the OECD classification of regions (Territorial Level 3) - which corresponds to, e.g., Statistical Divisions in Australia, Census Divisions in Canada, or Prefectures in Japan.

inventor-specific shocks which simultaneously increase an inventor's productivity in a given year t and make him co-invent with more high quality inventors in the previous year $t-1$.

Table 2: Main Results

	Benchmark				High Tech Sector	Broader Interactions Measures	
					Firm	Region	
High Quality Interactions (t-1)	0.0036***	0.017***	0.022***	0.026***	0.036***	0.000051***	0.0000017***
	(0.00057)	(0.0011)	(0.0013)	(0.0016)	(0.0042)	(0.0000016)	(0.000000078)
Low Quality Interactions (t-1)				-0.0037***			
				(0.00044)			
Age	0.000069	-0.032***					
	(0.00073)	(0.0011)					
Age^2	0.0000055	0.00049***					
	(0.000036)	(0.000051)					
Team Size	0.032***	0.026***	0.028***	0.028***	0.035***	0.028***	0.028***
	(0.00091)	(0.0012)	(0.0012)	(0.0011)	(0.0036)	(0.0012)	(0.0012)
Log Firm Size	-0.038***	-0.0086*	-0.036***	-0.035***	-0.065***		-0.037***
	(0.0028)	(0.0039)	(0.0038)	(0.0038)	(0.014)		(0.0038)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x Sector FE	Yes	No	Yes	Yes	Yes	Yes	Yes
Year x Region FE	Yes	No	Yes	Yes	Yes	Yes	Yes
Sector x Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Leader FE	No	Yes	Yes	Yes	Yes	Yes	Yes
N	1811493	1560713	1556569	1556569	285490	1556569	1556567
adj. R^2	0.15	0.18	0.19	0.19	0.13	0.19	0.19
F	287.9	425.2	307.6	229.3	58.3	786.7	335.9

Notes: All models include, in addition, sector and region fixed effects (not listed). Columns (1) and (3) - (7) also include year fixed effects. Columns (1)-(3) use the benchmark measure of interactions (past co-inventors). Column (4) also includes lower-quality past co-inventors. Column (5) limits the sample to the high-tech sector. Columns (6) and (7) consider, respectively, interactions measured as inventors in the same firm or in the same region. Standard errors clustered at the team leader are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Main results

Table 2 contains the results from the benchmark regressions. In this table, idea quality (the left-hand side variable) is always measured according to the benchmark measure, which is the forward citations of the patent in a 3-year window. In columns (1)-(3), interactions are measured in the benchmark way, as the total number of past higher quality co-inventors.

The effect of high quality interactions on patent quality is consistently positive and significant. With team-leader fixed effects (column (2) onwards), one additional high quality interaction increases subsequent patent quality by 0.02 standard deviations, which corresponds to a 0.06 level increase in citations per patent in a 3-year window.[31] What does this mean economically? If we use the mean patent quality for existing patents in the sample (i.e., mean citations per inventor in a 3-year window conditional on patenting) from Table 1, this represents a 4% increase. However, since inventors do not patent every year, a more accurate comparison would use the mean citations per

[31] In column (3) and beyond, the age coefficient drops from the regressions as we include both team leader and year fixed effects.

patent (in a 3 year window), per inventor, per year, in the sample (unconditional on patenting): One interaction represents a 21% increase in citations relative to that unconditional mean (see Table 1).

Table 3: Robustness Checks

	Repeated Interactions	Productivity Measures		Low Tech
		5-year cit.	Excl. Self-cit.	
High Quality Interactions (t-1)	0.010***	0.020***	0.020***	0.020***
	(0.00088)	(0.0012)	(0.0012)	(0.0013)
Team Size	0.028***	0.029***	0.019***	0.027***
	(0.0012)	(0.0011)	(0.0011)	(0.0011)
Log Firm Size	-0.036***	-0.047***	-0.029***	-0.031***
	(0.0038)	(0.0039)	(0.0037)	(0.0040)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Year × Sector FE	Yes	Yes	Yes	Yes
Year × Region FE	Yes	Yes	Yes	Yes
Sector × Region FE	Yes	Yes	Yes	Yes
Team Leader FE	Yes	Yes	Yes	Yes
N	1556569	1556569	1556569	1217811
adj. R^2	0.19	0.23	0.18	0.21
F	258.0	365.1	200.0	280.8

Notes: All models include in addition year, sector, and region fixed effects. Column (1) counts repeated interactions twice. Column (2) uses a 5-year window to measure citations (i.e., to measure idea quality); column (3) excludes self-citations. Column (4) limits the sample to low-tech sectors. Standard errors clustered at the team leader are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In column (4), we simultaneously include those past co-inventors who were more productive and those who were less productive than the inventor at the time of their encounter. If anything, the effect of high quality interactions increases to 30%. In the model, individuals benefit more from interacting with people better than they are, i.e., who have higher productivity.[32] In the data, low quality interactions, conditional on inventor fixed effects, have a very slight negative effect on productivity: They may be crowding out more productive interactions and wasting an inventor's time.

Column (3) restricts the sample to high-tech sectors according to the Eurostat classification.[33] In high tech sectors, the effect of one more high quality interaction is even stronger, equal to 8% of the sample mean conditional on patenting, and 39% unconditional on patenting.[34]

Columns (4) and (5) use the broader measures of interactions. Interactions as measured by the number of inventors in the firm (column (4)) and in the region (column (5)) who are more

[32] In fact, in the model, agents gain nothing by interacting with less productive people.

[33] The Eurostat classification uses specific sub-classes of the International Patent Classification (IPC) as defined in the trilateral statistical report of the EPO, JPO and USPTO. Macro high-tech technical fields are: Computer and automated business equipment; Microorganism and genetic engineering; Aviation; Communications technology; Semiconductors; Lasers. The list of sub-classes and their definition is provided by Eurostat at: `http://ec.europa.eu/eurostat/cache/metadata/Annexes/pat_esms_an2.pdf`

[34] See Table 3 for the effect in low-tech sectors, which is still strongly significant, but smaller in magnitude.

productive than the focal inventor, also significantly affect productivity. Naturally, the effects of these interactions are smaller than the effect of the benchmark measure, since the interaction with someone in the same firm or region is more distant and thus weaker. An inventor who has 10 more higher quality inventors in his firm has a 0.4% higher productivity. Additionally, an inventor with 10 more higher quality inventors in his region has a 0.3% higher productivity (both effects measured relative to the unconditional mean). However, these interactions can still be significant because of the sheer number of inventors in a firm or region (see Table 1). For instance, in the median region by inventor density, there are 506 inventors. In the top 25% regions, there are 2100 inventors; in the top 5% there are 13,200 inventors. Our estimates indicate that – even conditional on individual fixed effects – inventors who live in regions with many productive inventors, or work in firms who employ more productive inventors, are significantly more productive.

Robustness checks

To further test the robustness of the relation between interactions and productivity improvements, we consider additional measures of productivity and interactions in Table 3. In column (1), we count repeated interactions: If inventor i interacts with inventor j in two different years, that interaction is counted twice. In column (2), we use a five-year window to count forward citations. In column (3), we exclude self-citations. In column (4), we focus only on low tech sectors. In all cases, the effect of interactions is significant and the magnitudes are very robust.

We also re-estimate the model using USPTO data.[35] While we prefer EPO data for our benchmark estimation due to its representativeness for international comparisons (Miguelez et al. 2013) and higher average quality (De Rassenfosse et al. 2021), Table A-2 reproduces our main results with USPTO data, yielding similar outcomes. High-quality interactions are strongly associated with future patent quality, with the largest effects at the research team level and positive effects at the firm and region levels.

Empirical Variation: Switching Between Leading and Accumulating Interactions

Including team leader fixed effects in our regressions means that the direct variation affecting the coefficient of lagged high-quality interactions comes from *switcher* inventors–those who alternate between leading patents and accumulating interactions as team members. We study these switcher inventors' careers before their first patents as team leaders and examine how learning environments in different firms and regions influence their learning opportunities and interactions.

Switcher inventors represent 24% of unique team leaders and 40% of observations in the benchmark regression sample. Appendix Table A-3 provides summary statistics on these inventors. Before becoming team leaders, switchers have fewer interactions than other team leaders but work in larger companies and regions. Over 75% remain with the same firm, and more than 90% stay in the same region where they first patented. After becoming team leaders, switchers average 1.4 role switches and lead 8.5 patents, with most interactions occurring within the same firm and region.

[35] We use disambiguated inventor data from PatentsView.

Table 4: High Quality Interactions – Contemporaneous Correlations Switchers

	High Quality Interactions				
Age	0.402***	0.386***	0.403***	0.388***	0.379***
	(0.0129)	(0.0126)	(0.0130)	(0.0127)	(0.0146)
Age^2	-0.0111***	-0.0101***	-0.0111***	-0.0103***	-0.0101***
	(0.000845)	(0.000827)	(0.000844)	(0.000829)	(0.000873)
Log Firm Size		0.517***		0.534***	0.569***
		(0.0116)		(0.0121)	(0.0133)
Log Region Size			-0.0228	-0.113***	-0.113***
			(0.0158)	(0.0156)	(0.0173)
Firm Switching (dummy)					0.00687
					(0.0394)
Team Leader FE	No	No	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes
N	618034	618034	613956	613956	541792
adj. R^2	0.177	0.228	0.177	0.230	0.217
F	270.6	327.8	243.8	302.2	284.3

Note: Contemporaneous correlations of high-quality interactions with age, firm size, region size, and firm switching. Standard errors clustered at the team leader are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

To further study the factors associated with the accumulation of high-quality interactions, in Table 4, we show the contemporaneous correlations of high-quality interactions with age, firm size, region size, and firm switching. Older team leaders have more interactions, likely due to more collaboration opportunities. Inventors in larger firms have higher-quality interactions, with firm size accounting for most variation, while region size is negatively correlated when considered simultaneously. Firm switching is not significant, and its contribution to variance is modest, as shown by the adjusted R^2 statistics.

Finally, we study the heterogeneous effects of interactions by splitting the sample of team leaders based on their switching behavior across roles (team leader $\rightarrow$ team member), across firms, and across regions. Specifically, we modify our benchmark specification (14) by including lagged high-quality interactions interacted with dummies for inventors switching frequencies of 1, 2, 3, and 4 or more times,[36]

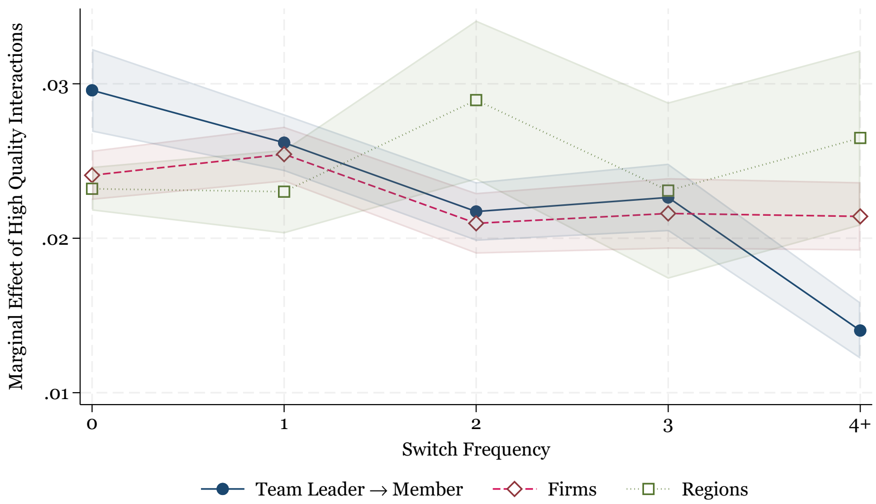
$$q_{j,t} = \beta_1^0 \text{Interactions}_{i[j],t-1} + \sum_{k=1}^{4^+} \beta_1^k \text{ Interactions}_{i[j],t-1} \times \mathbb{I}\left[Switch_{i,t} = k\right] + \text{Controls} + \text{Fixed-Effects} + \varepsilon_{j,t}.$$

Figure 4 shows how the marginal effect $(\beta_1^0 + \beta_1^k)$ of the lagged interactions on patent quality changes

[36] We include the same controls and fixed-effects as in regression (14) and add the switching dummies $\mathbb{I}\left[Switches_{i,t} = k\right]$ as regressors.

with respect to the switching frequency. The solid blue line indicates that, for inventors who switch roles more frequently, the positive impact of past high-quality interactions on patent quality decreases. This suggests that the benefits of collaborating with high-quality co-inventors decrease with frequent role switching, likely due to reduced learning opportunities. In contrast, the maroon dashed line (switches across firms) and the green dotted line (switches across regions) show a relatively stable effect, indicating that valuable learning opportunities persist when inventors switch to new firms or regions.

Figure 4: Marginal Effects of High-Quality Interactions by Switching Frequency



Notes: Marginal effects from running a regression of lagged high-quality interactions × dummies of switching frequency 1,2,3,4+. The solid blue line with dots denotes switches across roles (team leader → team member), the maroon dotted line with diamonds switches across firms, and the dotted green line with squares switches across regions. Shaded regions correspond to the 95% confidence intervals. All regressions have the same controls and fixed effects as the benchmark specification.

4 Estimation

In this section, we bring together the model from Section 2 and the empirical results from Section 3 to estimate the model using the new data on productivity, innovation, and interactions. We explain our numerical algorithm, the parameters estimated, and the moments that we target. We then show that the model fits the data very closely, not only for targeted moments but also for non-targeted ones.

4.1 Computational Algorithm, Functional Forms and Parameters

We start by describing our numerical algorithm. For any given set of parameters, we simulate the model for $n_{sim} = 100,000$ individuals for $T = 25$ years.[37] We initialize the simulations with a given

[37] We discretize each year with a size step of $\Delta t = 0.2$.

initial cross-sectional distribution $F(z,0)$ (see Table 5).

Table 5: Functional Forms

Function	Description
$q(z) = z^{1-\eta} n^{\eta}$	Idea production function.
$c(m,s) = \frac{\kappa(s)}{2} m^2$	Interactions cost function.
$F(z,0) = \frac{1}{1+\lambda z^{-1/\theta}}$	Initial cross-sectional productivity distribution.
$\Gamma(x,0) = \frac{1}{1+k_0 x^{-1/\theta}}$	Age zero productivity distribution.
$\Psi(x,s) = \frac{1}{1+\lambda_e(1+s^{\nu})x^{-1/\theta}}$	External (age-dependent) learning distribution.

In each period, there is a constant probability δ that an individual survives to the next period. We can compute the number of successful interactions (i.e., the number of times an agent improves his productivity by learning from someone better than him) and the number of successful draws from the external distribution. We thus obtain the realized productivity distribution $F(z,t)$ at each time t, as well as the age-conditional productivity distributions, $G(z,s,t)$. We then use the distribution $F(z,t)$ to compute the paths of researchers' wages $w(t)$ and productivity cutoffs $\bar{z}(t)$ in the occupational choice problem. In the BGP, the economy reaches the invariant distributions $\Phi(z)$ and $\Gamma(z,s)$. At age zero, new inventors draw their productivity from the distribution $\Gamma(z,0)$.[38] When an individual dies, he is replaced by another individual with a productivity draw from the same initial distribution $\Gamma(z,0)$.

Using the simulated model, we compute the values of key data moments described below and iterate on the parameters until we find the ones for which the model-generated moments best match the data moments. We match the moments in the data when the economy is on a BGP. To parameterize the model, we suppose that the initial cross-sectional distribution $F(z,0)$, the age zero distribution $\Gamma(x,0)$, and the external source distribution $\Psi(x,s)$ are log-logistic. In Table 5, we summarize all the functional forms relevant to the calibration of the model.

We estimate twelve parameters of the model – listed in Table 6 – and captured by the vector:

$$\chi := (m_X, \lambda, \theta, \nu, \eta, \delta, k_0/k, \alpha, \iota, \kappa_0, \kappa_1, \kappa_{min}).$$

4.2 Data Moments and Estimation

We estimate the twelve-parameter vector χ by matching key data moments. We target the following ten groups of moments for a total of 87 moments:

1. Average number of high-quality interactions of team leaders by age (estimated at 18 age values, each corresponding to one moment).
2. Average idea (patent) quality of team leaders by age (estimated at 18 age values).

[38] In the data, age zero corresponds to the first year in the sample, i.e., the first year in which an inventor patents.

Table 6: Estimated Parameters

Parameter	Description	Value
m_X	Rate of draws from external learning source.	0.524
λ	Location parameter of initial distribution $F(z,0)$.	1.160
θ	Tail parameter of productivity distributions $F(z,0)$, $\Gamma(x,0)$ & $\Psi(x,s)$.	0.271
ν	Exponent on location parameter $\rho(s) = \lambda\left(1+s^{\nu}\right)$ of $\Psi(x,s)$.	1.136
η	Team leader's span of control.	0.048
δ	Death hazard rate.	0.319
$\frac{k_0}{k}$	Relative location parameter of age zero distribution $\Gamma(x,0)$.	0.240
α	Final good elasticity with respect to aggregate productivity.	0.121
ι	Fraction of inventors with new research ideas.	0.610
κ_0	Interactions Costs: constant term.	0.601
κ_1	Interactions Costs: linear term.	-0.032
κ_{min}	Interactions Costs: minimum value.	0.110

3. Average productivity of all inventors by age (estimated at 18 age values).
4. Team size distribution (discretized and estimated at 20 team size values).
5. The age distribution (discretized and estimated at 18 age values).
6. The share of inventors who are team leaders.
7. The marginal effect of high-quality interactions on idea quality.
8. The marginal effect of age on idea quality.
9. The marginal effect of age^2 on idea quality.
10. The economy's growth rate.

Some of these groups of moments are life cycle profiles (1-3) or entire distributions (4 and 5), which are matched at discretized values. Table 7 summarizes the moments we target.

Two of the moments are the marginal effects of, respectively, age and interactions on productivity (moments 7 and 8). Intuitively, conditional on the number of interactions and all our other fixed effects and controls (including also individual fixed effects), the effect of age on productivity helps calibrate the strength of the external learning channel since it captures improvements in productivity that arise over time but do not directly come through interactions. The effects of age and interactions were estimated in the regressions in Section 3 with many different measures of productivity and interactions, as well as with several different specifications. We run an equivalent regression using data generated from the simulated model:

$$\text{Productivity}_{it} = \beta_0 + \beta_1 \text{High-quality interactions}_{it-1} + \beta_2 \text{Age}_{it} + \beta_3 \text{Age}_{it}^2 + \varepsilon_{it} \tag{15}$$

where $i \in \{1, ..., n_{sim}\}$ indexes simulated inventors and $t = \{1, ..., 25\}$ age. We then match the marginal effect of interactions β_1 and age β_2 from the actual and moment simulated, data.[39]

Table 7: Targeted Moments

Moment	Weight ω_i	Description	Model	Data
1-18	$\frac{1}{10}\frac{1}{18}$	Average high-quality interactions of team leaders by age	see Fig. (5a)	see Fig. (5a)
19-36	$\frac{1}{10}\frac{1}{18}$	Average Idea quality of team leaders by age	see Fig. (5b)	see Fig. (5b)
37-54	$\frac{1}{10}\frac{1}{18}$	Productivity of all inventors by age	see Fig. (5c)	see Fig. (5c)
55-74	$\frac{1}{10}\frac{1}{20}$	Team size distribution	see Fig. (5d)	see Fig. (5d)
75-82	$\frac{1}{10}\frac{1}{18}$	Age distribution of team leaders	see Fig. (5e)	see Fig. (5e)
83	$\frac{1}{10}$	Fraction of team leaders	0.501	0.531
84	$\frac{1}{10}$	Regression coefficient on high-quality interactions, β_1	0.061	0.062
85	$\frac{1}{10}$	Regression coefficient on age, β_2	0.039	0.040
86	$\frac{1}{10}$	Regression coefficient on age^2, β_3	-0.001	-0.001
87	$\frac{1}{10}$	Growth rate	0.025	0.025

We choose the values for all the estimated parameters in Table 6 that minimize the weighted sum $\mathcal{L}(\chi)$ of the absolute difference between the model-implied moments and the data moments:

$$\mathcal{L}(\chi) := \min \sum_{i=1}^{87} \omega_i \frac{|\text{model}(i) - \text{data}(i)|}{\frac{1}{2}|\text{model}(i)| + \frac{1}{2}|\text{data}(i)|},$$

where ω_i is the weight assigned to moment i, reported in Table 7. We assign the weights ω_i so that all nine groups of moments (each represented by a row in Table 7) receive equal weight. Thus, moments within a group receive a weight of $1/10$ multiplied by $1/n_p$ where n_p is the number of discretized points of the distribution that we match.

Parameter Estimates

Table 6 reports the estimates we obtain from our calibration procedure. The average meeting cost parameter is equal to $\int_0^\infty \kappa(s) d\Omega(s) = 0.50$. This leads to a average endogenous interaction rate of $\int \int_0^\infty m^*(x,s) d\Omega(s) d\Phi(x) = 1.18$, larger than the external source meeting rate $m_X = 0.52$. The location parameter of the initial distribution is equal to $\lambda = 1.16$, and the tail parameter of the distributions is $\theta = 0.27$. Parameter ν determines the concavity in age of the age-dependent location

[39] For the data regression, we include a time trend, sector fixed effects, and country fixed effects, which need to be filtered out to make the setting comparable to the model. Note the coefficients differ from the ones in Table 2 as we do not include the other fixed effects.

parameter of the external distribution $\varrho(s) = \lambda(1 + s^{\nu})$. A larger value of ν means that the average productivity of external ideas grows faster during the first years of experience. In our calibration, $\nu = 1.14$. The relative importance of team members in patent production is $\eta = 0.05$, which implies that the team leader's span of control is equal to $(1 - \eta) = 0.95$. The death hazard rate is equal to $\delta = 0.32$. The elasticity of final good production to productivity is equal to $\alpha = 0.12$.[40] Finally, ι reflects the fraction of inventors that get new research ideas each period.[41]

4.3 Goodness of Fit

Table 7 reports the target for the moments from the data and the values we obtain in the estimated model. In cases in which the moment is a distribution function, we do not report a value but instead, provide the reference to the relevant figure that shows how the distribution generated from the simulated model matches the one in the data.

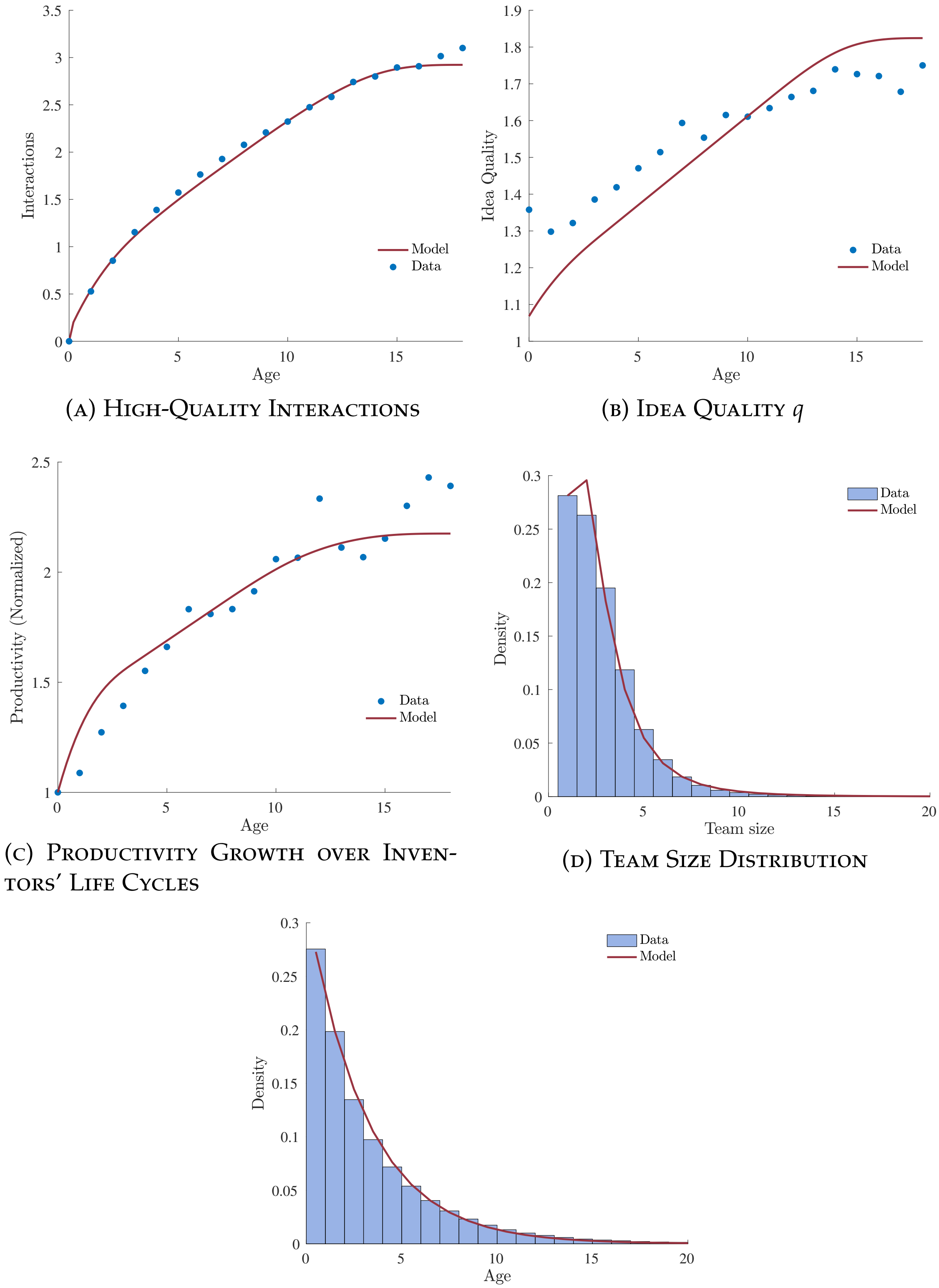
Our calibration closely fits all the targeted moments; we are able to capture most of the salient and peculiar features of the data. This can be seen most vividly in Figure 5. Panel (A) shows that the model implies that the average number of high quality interactions for team leaders is increasing in age and that the relation is close to linear; this is also exactly what we see in the data. The model also predicts that the average idea quality produced by team leaders and their teams is increasing and concave in age, which is also true of the data (see panel (B)). The productivity of all inventors is increasing and concave in age in both the model and the data (see panel (C)). Team leaders account for 53% of individuals, and the team size distribution is decreasing (see panel (D) and Table 7). The age distribution is also decreasing, with near exponential decay (see panel (E)).

Our estimated parameters jointly determine the value of these moments. To better understand the variation that identifies these parameters, we study the elasticity of the moments with respect to changes in the parameters around their estimated value (see Table A-1). Some of the moments are affected only by a few parameters, such as *the age distribution* that is almost entirely determined by the death rate δ, and *the average team size*, which responds mainly to the Pareto tail parameter θ. Other moments have more complex dependencies on the parameters that are nonetheless telling about the underlying mechanisms of the model. For instance, the interaction related moments (*average interactions* and *the interactions coefficient* β_1) are not only affected by the cost parameters κ's (via the endogenous interaction rate) but also by the learning opportunities reflected in how fat is the tail of the distribution θ and the team structure determined by the fraction of individuals that get new ideas ι. In turn, the inventors' productivity moments respond in levels (*idea quality*) to the location parameter λ and the fatness of the tail θ, and in their life-cycle evolution (*inventors' productivity*) to more frequent exogenous meetings and when the tail parameter increases in response to more learning opportunities as they age. The age coefficient β_2 increases when inventors die more frequently, and the fraction of team leaders decreases in the tail parameter and in the relative importance of team members in the idea production η.

[40] Intuitively, this parameter is calibrated to match the empirical growth rate.

[41] We do an additional normalization to make the units of productivity (log of citations) comparable to the number of inventors in the data. To do so, we calibrate a multiplicative parameter so that the average level of the team size distribution matches.

Figure 5: Targeted Moments – Fit between model and data



(A) High-Quality Interactions

(B) Idea Quality q

(C) Productivity Growth over Inventors' Life Cycles

(D) Team Size Distribution

(E) Age Distribution

Notes: Panel (A) represents the high-quality interactions of a team leader as a function of his age (age 0 is the first year in the patent data). Panel (B): represents the de-trended idea quality $q(z,t)e^{-g_q t} = q(z,t)e^{-g(1-\eta)t}$, where $g_q = g(1-\eta)$ is the growth rate of idea quality, as a function of team leader's age. Panel (C) depicts average inventor productivity z across all inventors at a given age. It is normalized by productivity at age 0. Panel (D) shows the distribution of the team size n, and Panel (E) shows the distribution of age (time in the sample).

Endogenous Meetings

Any factor that influences the costs or benefits of interactions will affect the endogenous meeting rate. To analyze how the meeting rate changes in relation to the parameters, we estimate the Jacobian matrix (outlined in Appendix A-2), which measures the marginal responses of endogenous meetings with respect to all the parameters in the model. We study the marginal response of the average search efforts $m^* = \int \int m^*(x,s) d\Omega(s) d\Phi(x)$ to changes in the model's parameters around their estimated value.[42] Intuitively, the average endogenous meeting rate is strongly decreasing in the cost parameter κ_0 that determines the initial level of the cost function. On the benefit side, we find that a fatter tail parameter θ enhances learning opportunities, leading to an increase in the endogenous meeting rate. Conversely, a higher death rate δ reduces the incentives to search as individuals have less time on average to reap the benefits of learning. Additionally, there is a substitution effect between the two types of learning: an increase in the external source meeting rate m_X partially offsets the incentives to search.

4.4 Fit for Non Targeted Moments

Idea quality Pareto tail

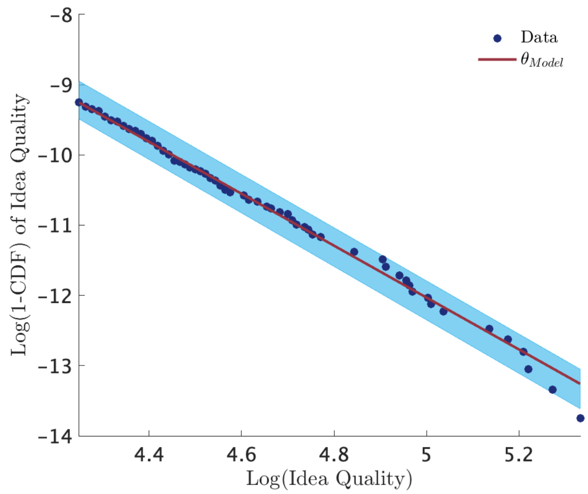
We now check the goodness-of-fit of the estimated model for some non-targeted moments, namely the tail of the idea quality distribution. As we discuss above, the tail of distribution is fundamental in determining growth and the learning opportunities of individuals. Figure 6 depicts the log counter cumulative distribution ($\log(1 - CDF)$) for log idea qualities above the 99.9th percentile.[43] The tail of the distribution generated by our estimated model strongly resembles the one observed in the data. The estimated tail parameter in the model $\theta = 0.271$ is within the 95% confidence interval of the tail parameter, $[0.253, 0.274]$, estimated using data from the tail of the distribution (from the 99.9th to the 99.9999th percentile).[44]

[42] These elasticities correspond to the first row in the Jacobian Matrix A-1.

[43] The 99.9th percentile corresponds to patents with more than 70 citations in the 3-year window.

[44] We repeat this same exercise for the US and Germany. Figures are presented in Appendix A-3.

FIGURE 6: TAIL OF IDEA QUALITY DISTRIBUTION



Note: The figure plots the log of the counter CDF (log(1-CDF)) of the log idea quality $log(q)$. The red solid line is a linear relation predicted by the model estimated θ anchoring the initial point to the data. The shaded area is the 95% confidence interval of a regression using data between the 99.9th and the 99.9999th percentile.

Earnings and interactions

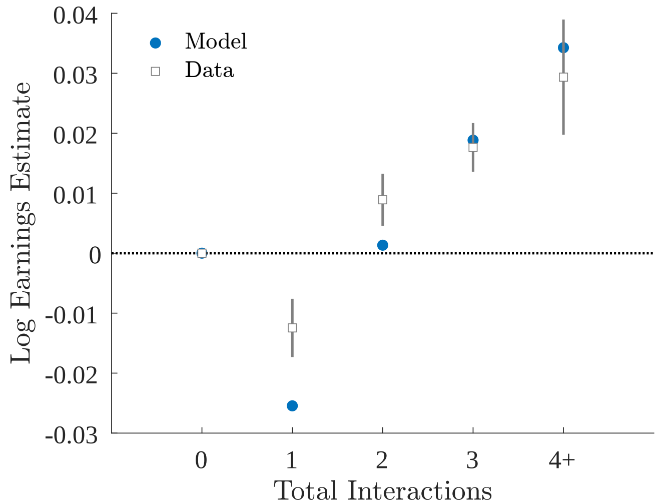
One key prediction of the model is the relation between inventors' monetary returns to innovation and their past interactions. Using a novel dataset from Akcigit and Goldschlag (2025) combining US Census and inventor data, we compare the marginal effects of the number of interactions on inventor earnings [45],

$$\log earn(x_i) = \alpha_0 + \sum_j \zeta_j Interactions[j] + Age_i + \epsilon_i.$$

Figure 7 shows that the marginal effects in both the model and the data are closely aligned. Notably, financial returns initially decline before increasing as the number of interactions rises, with the model's results closely matching the estimated coefficients. In the model, the initial decrease in returns can be attributed to the selection of individuals who never interact, as they tend to have, on average, higher productivity than those who have had one interaction when also controlling for age.

[45] In the model, we focus on the earnings of team leaders who are the residual claimants of the innovations.

Figure 7: Untargeted Moments: Earnings and Interactions



Note: The figure plots the marginal effects of the number of interactions on the logarithm of earnings both in the model and in data, using the data from Akcigit and Goldschlag (2025).

5 Quantitative Analysis and Policy Experiments

As shown in Sections 4.3 and 4.4, the estimated model fits the data very well for both targeted and non-targeted moments. The model is complete enough to be used to understand past and contemporaneous changes in knowledge diffusion, productivity, and innovation. To illustrate this, we now put our estimated model to work and quantitatively explore four policy experiments:

1. **Importance of interactions**: We quantify the importance of interactions for productivity and growth, relative to other learning channels.

2. **Reducing interaction costs:** What happens when interaction costs are reduced? Concretely, this can be interpreted as companies or industries adopting a "Google model," with open-floor offices and designated hang-out times that foster interactions. It also can be mapped to the rise of IT which facilitates interactions.

3. **Reduced access to external ideas:** We study the effect of reduced access to external ideas. This highlights potential downsides of agglomeration and the "paradox of proximity."

4. **U.S. vs. Germany:** We estimate the model separately for two large innovating countries, Germany and the U.S., to compare their innovation production functions and team dynamics. What would happen to U.S. growth were it to inherit the (lower) interaction costs of Germany?

There are many other possible counterfactual analyses that can fruitfully be done using our estimated model. Moreover, the data can be sliced in several other ways: e.g., the model could be separately estimated by technological field or time period to see how the innovation production and research teams dynamics differ.

5.1 Interactions versus External Learning

We start by computing the contributions of interactions to economic growth.[46] For this exercise, we shut off, in turn, the endogenous interaction channel (setting $m(x, s) = 0 \ \forall x, s$) and the exogenous learning channel (setting $m_X = 0$). Table 8 shows the resulting growth rates in these two cases.

Table 8: The Role of Interactions and External Learning

	Baseline	No interactions ($m(x, s) = 0$)	No external learning ($m_X = 0$)
Fixed $m(x, s)$	2.50%	0.24%	1.51%
Endogenous $m(x, s)$	2.50%	0.24%	1.87%

Notes: This table estimates the growth rate under the benchmark parameters when shutting down each of the learning channels. In the first row we estimate the growth the rate for each case while fixing the endogenous meeting rate to the one obtained in the baseline case. In the second row we estimate the growth by allowing agents to internalize the change in the parameters when optimizing $m(s)$.

There are three key findings.

1. Shutting down any of the two learning channels significantly reduces growth.
2. The endogenous interaction channel contributes more to growth than the exogenous learning channel.
3. There is a strong complementarity between the two learning channels.

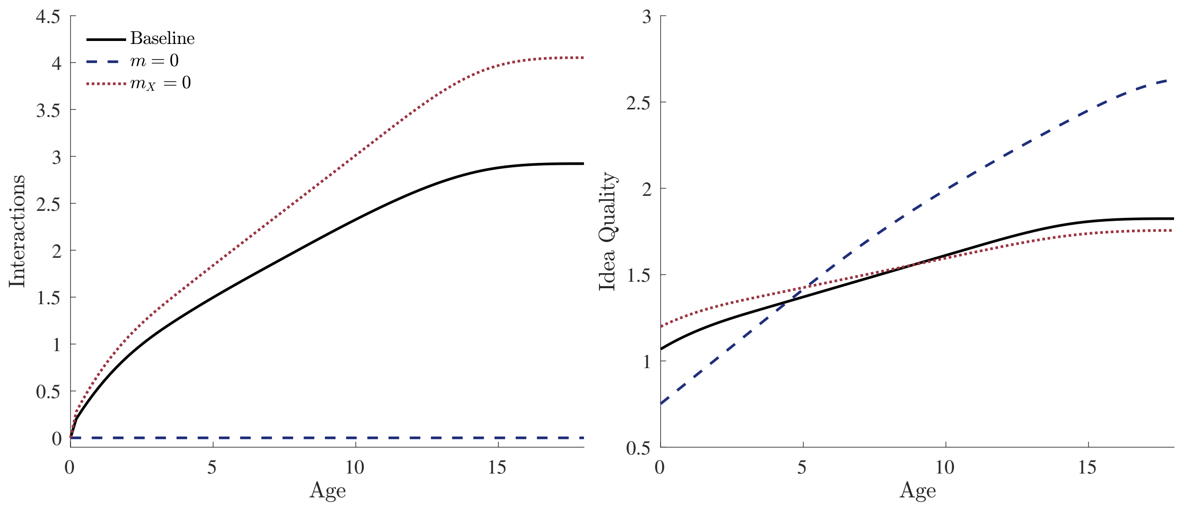
Clearly, both channels significantly contribute to growth; when either one is shut off, growth is reduced by at least 40%. Nevertheless, when the endogenous interaction channel is shut down (case with $m(s) = 0 \ \forall s$), growth is reduced to a striking 0.2%. If the external learning channel is shut down (case with $m_X = 0$), the growth rate drops by less, to 1.5%. In fact, recall from Section 2 that in the special case in which the external productivity distribution $E(z, s, t)$ is not growing, the growth rate is $m\theta$. In this case, shutting down interactions reduces growth to 0 (but, conversely, shutting down external learning would still lead to positive growth). This case, akin to the one in Lucas and Moll (2014), emphasizes even more starkly our quantitative finding that interactions are vital for growth.

The third finding is that the contribution of each channel individually adds up to much less than the total growth rate when both channels are active. This complementarity between the two learning channels is key: when agents learn more from the external source, each endogenous interaction becomes more valuable in expectation as one can hope to acquire that new external knowledge indirectly by meeting others. Individuals then optimally choose a higher endogenous meeting rate. The second row of Table 8 explicitly shows this complementarity as growth drops by less, to 1.9%, when individuals choose the endogenous meeting rate.

Figure 8 shows the effect of shutting down the endogenous and exogenous channels on the number of interactions and on idea quality. The solid black curve represents the estimated benchmark

[46] For this, it is useful to recall the discussion in Section 2, which described in what ways the growth rate depends on the rate of external learning and endogenous interactions and some special cases.

Figure 8: The Role of Interactions



(A) High-Quality Interactions (B) Idea Quality q

Panel (A) represents the high-quality interactions of a team leader as a function of his age (age 0 is the first year in the patent data). Panel (B): represents quality $q(z,t)$ as a function of the team leader's age. In this panel, the right vertical axis is for the baseline scenario. In each panel, the solid black line represents the baseline case, with both the endogenous interaction channel and the external learning channel active. The blue dashed line corresponds to the case with no interactions within teams ($m(x,s) = 0$). The red dotted line corresponds to the case with no external learning ($m_X = 0$).

case from Section 4 (i.e., the case in which the two channels are active).[47] Panel (A) highlights that when the exogenous learning rate is zero ($m_X = 0$), team leaders have more interactions than in the benchmark case. Remember, individuals interact only while they are team members, so more interactions by age means it takes more time for individuals to become team leaders. Panel (B) shows that the life cycle of de-trended idea quality is starkly different for each case. In the absence of team member interactions ($m(x,s) = 0$) learning is slower and more spread out throughout the inventors' lives. In contras,t when the external learning is shut off, all the improvements in productivity happen early with limited gains before it flattens, which illustrates the importance of an external pool of ideas to keep up learning opportunities. Put simply, with no external sources of ideas, a group of individuals will eventually learn most (albeit not all) there is to be learned from each other and innovation will slow down.

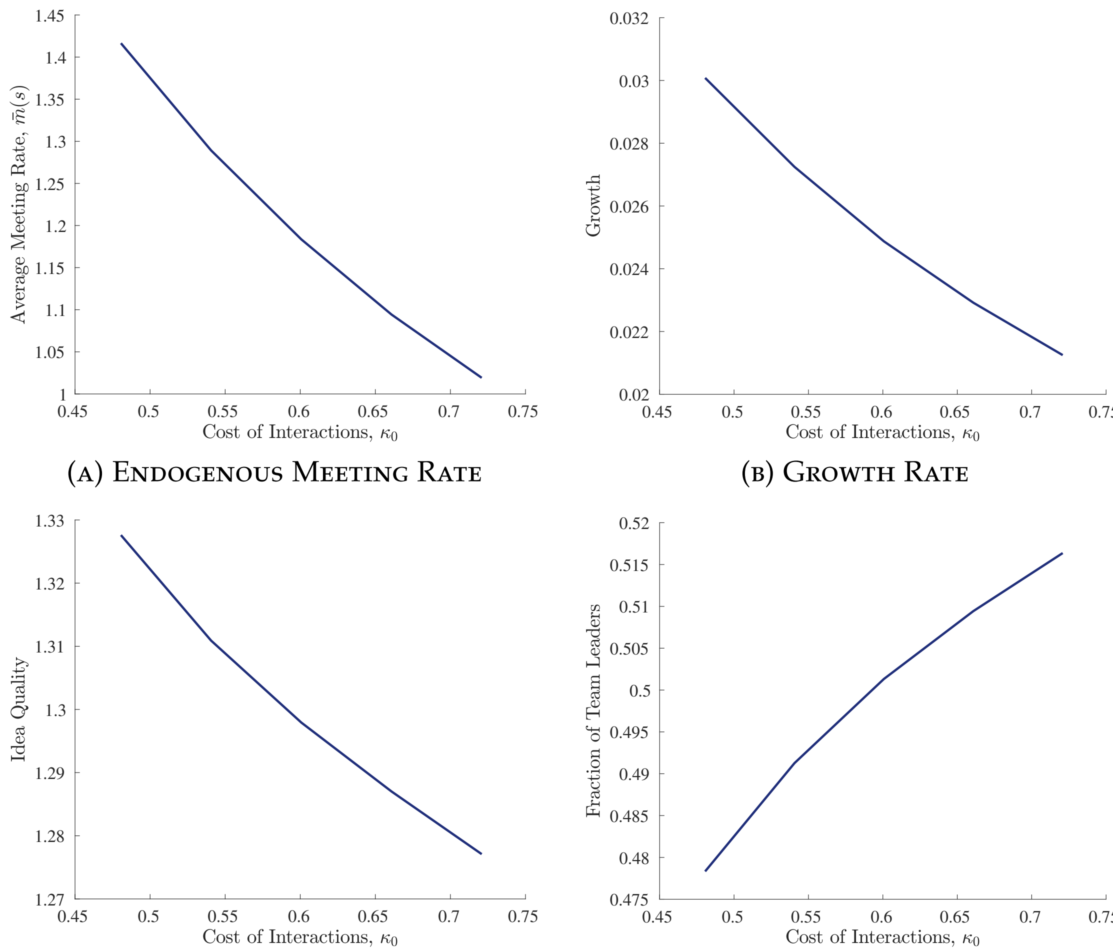
5.2 Lowering Interaction Costs: Information Technology or the "Google Model"

What happens if the cost of interactions, $\kappa(s)$, decreases? This scenario can mimic the spread of information technologies (IT) that make communication and, hence, interactions easier, even across larger geographical distances. IT can also break down language barriers (e.g., through real-time translation or easy tools such as Google translate) or matching frictions (through online platforms and specialized forums). Many dedicated programs and apps have appeared to make interacting

[47]In the Appendix A-6 we also show the effect on other moments.

easier and cheaper, e.g., Zoom, Teams, Whatsapp, or FaceTime.[48] A decrease in interaction costs can also mimic the rise of new business models, such as the "Google model" that allows for designated spaces and times for interactions or, more generally, the open space office.

Figure 9: The Effects of Interaction Costs κ_0



(a) Endogenous Meeting Rate

(b) Growth Rate

(c) Idea Quality

(d) Fraction of Team Leaders

Notes: Panel (A) represents the idea quality $q(z,t)$. Panel (B) depicts the yearly growth rate of product $g_y = \alpha(1-\eta)g$. Panel (C) presents the idea quality and Panel (D) the fraction of team leaders. All plots as functions of interaction intercept from the cost of interactions function κ_0.

A lower interaction cost κ_0 will, all else equal, increase the endogenous average meeting rate $m^* = \int_0^\infty m^*(x,s)d\Omega(s)$. As researchers have an easier time meeting others, they learn more, and the quality of innovation and the growth rate are higher. These results lend support to the many attempts in the real world to make interactions easier through conferences, socializing events, and common spaces.

A perhaps less immediate result is that lower interaction costs decrease the fraction of team leaders. The decrease in κ_0 increases the cutoff $\hat{x}(s)$ above which individuals become team leaders

[48]On the differential impact of I.T. in Europe and in the U.S., see Bloom, Sadun, and Reenen (2012).

(see Figure A-4a). At the same time, the productivity distribution slightly improves, creating more mass to the right of the cutoff (see Figure A-4b). On balance, there are more team leaders (see Figure 9d), and thus more competition for skilled researchers hired as team members. Therefore, teams become smaller on average and, due to the entering of lower quality team leaders, the average idea quality decreases (see Figure 9c).[49]

5.3 Limiting Access to External Knowledge

What happens if access to external learning sources is reduced? There are many concrete situations in which individuals become less exposed to external sources of knowledge. First, there may be a strong agglomeration and geographical concentration of talent in some areas, which, paradoxically, may lead to a great deal of interactions with similarly-minded people, but little inflow of new knowledge. This is a version of the "proximity paradox," whereby too much cognitive or geographical proximity with the same group of people – importantly, without additional external inflow of new knowledge – can hinder innovation. Second, there may be very intensive specialization that makes it hard to understand or adopt external knowledge (even if interactions with others from the same specialization are very productive, as everybody speaks the same "language.") Third, protectionism, with an ensuing reduction in trade and foreign direct investment (FDI), can lead to lower inflow of new knowledge.[50] In all these settings, each specialized or concentrated group of researchers becomes "an island" that gets draws from the outside world's knowledge only rarely; almost all of the learning happens from interacting with other people.

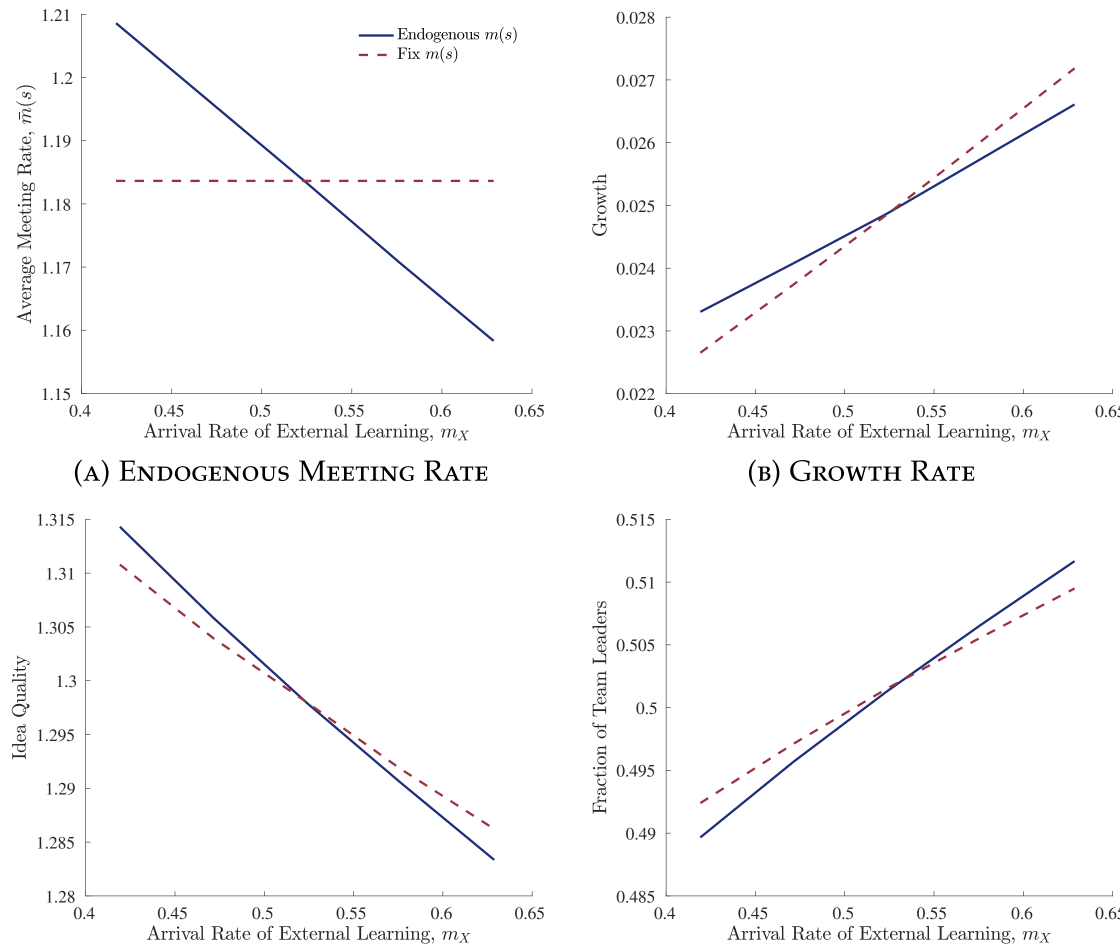
Figure 10 shows the effect of access to external knowledge when $m(x, s)$ is chosen endogenously by agents (solid blue line) and when it is held fixed at the implied benchmark calibration value with $m_X = 0.524$ (dotted red line). The figure shows that if individuals have less access to external knowledge (m_X decreases), the productivity of individuals grows slower, with a lower elasticity in the endogenous case as individuals optimally increase the rate of interactions. More access to external knowledge increases the fraction of team leaders, decreasing the average team size. This implies a decrease in average quality of idea as the team leader productivity thresholds decrease. See also Appendix Figure A-5 for the full distributions of productivity and cutoffs for different rates of external learning.

Historically, there have been many examples of external knowledge channels improving knowledge and productivity in previously autarkic societies. Roads, railroads, sea routes, transport infrastructure and communication technologies were instrumental in facilitating knowledge "from the outside" and, subsequently, speeding up innovation. The economic history papers presented in the introduction describe many such episodes (see also Agrawal, Galasso, and Oettl (2017)).

[49]See also Appendix Figure A-4, which shows the effect of κ_0 (for values $\kappa_0 \in \{0.54, 0.60, 0.66\}$) on the cutoffs and the distribution.

[50]See the papers on trade and technology diffusion discussed in Section 1.1, keeping in mind that in our case, it is *knowledge* used to produce innovation, and not directly production technologies that are diffused.

Figure 10: Effect of Access to External Learning and Ideas



(a) Endogenous Meeting Rate

(b) Growth Rate

(c) Idea Quality

(d) Fraction of Team Leaders

Notes: Panel (A): shows the endogenous meeting rate. Panel (B) depicts the yearly growth rate of product $g_y = \alpha(1-\eta)g$. Panel (C) represents the idea quality $q(z,t)$. Panel (D) is the fraction of team leaders. All plots as functions of external ideas meeting rate m_X. The solid blue line represents the case where $m(x,s)$ is endogenously updated, and the dotted red line when $m(x,s)$ is fixed at its value in the benchmark calibration with $m_X = 0.524$.

5.4 The U.S. versus Germany: The Effects of Interactions in Different Labor Markets

We can also use our model to understand how interactions affect productivity and growth in different economic environments. It is possible to slice the data in many different ways, and to estimate the model separately by time period, by technology sector, or by country. Here, we estimate our model separately for the U.S. and Germany – two major patenting countries with very different labor markets and education (i.e., human capital acquisition) systems. Related to this, these two countries have very different life cycle profiles of interactions and productivity (as illustrated in Figure 11). Using the same estimation procedure outlined above, but applied to the data from each country separately, we obtain the country-specific parameters summarized in Table 9. The model fits each country's data remarkably well. The moments in the data and model are summarized in

Table 10 and Figure 11.[51]

Table 9: Estimated Parameters for the U.S. and Germany

	Full Sample	U.S.	Germany
m_X	0.524	0.743	0.338
λ	1.160	5.753	0.100
θ	0.271	0.328	0.280
ν	1.136	0.522	1.669
η	0.048	0.014	0.128
δ	0.319	0.310	0.236
$\frac{k_0}{k}$	0.240	0.462	0.080
α	0.121	0.074	0.137
ι	0.610	0.524	0.575
κ_0	0.601	1.055	0.488
κ_1	-0.032	-0.012	-0.039
κ_{min}	0.110	0.093	0.075

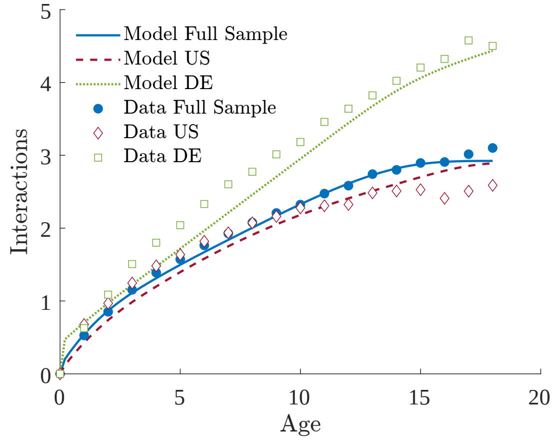
Table 10: Moments for the U.S. and Germany

	Model Full Sample	Data Full Sample	Model US	Data US	Model DE	Data DE
Interactions	Figure 11a	Figure 11a	Figure 11a	Figure 11a	Figure 11a	Figure 11a
Patent Quality	Figure 11b	Figure 11b	Figure 11b	Figure 11b	Figure 11b	Figure 11b
Productivity All	Figure 11c	Figure 11c	Figure 11c	Figure 11c	Figure 11c	Figure 11c
Team Size Distribution	Figure 11d	Figure 11d	Figure 11d	Figure 11d	Figure 11d	Figure 11d
Age Distribution	Figure 11e	Figure 11e	Figure 11e	Figure 11e	Figure 11e	Figure 11e
Interactions Coefficient, β_1	0.061	0.062	0.082	0.086	0.032	0.034
Age Coefficient, β_2	0.039	0.040	0.054	0.060	0.014	0.014
Age Coefficient^2, β_3	-0.001	-0.001	-0.002	-0.002	0.001	0.000
Fraction of Team Leaders	0.501	0.531	0.451	0.499	0.392	0.538
Growth Rate	2.5%	2.5%	2.74%	2.74%	2.42%	2.42%

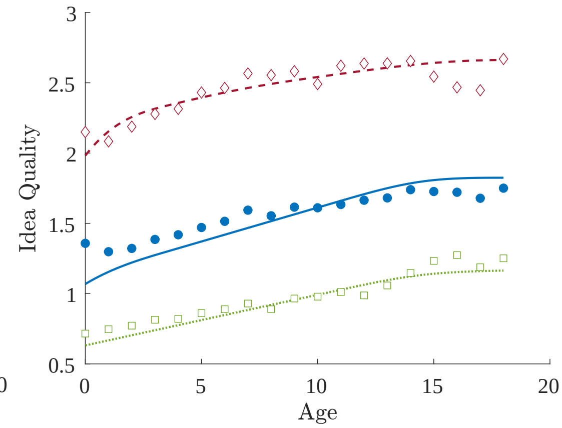
Panel (A) of Figure 11 shows that team leaders in Germany have, on average, more high-quality interaction than those in the U.S. The endogenous meeting rate is higher for Germany than for the U.S. (9% higher on average). This is a consequence of both lower meeting costs in Germany ($\bar{\kappa} = 0.34$ for Germany versus $\bar{\kappa} = 1.01$ for the U.S.) and of a higher expected benefit of each meeting. Team leaders in Germany start with ideas of lower quality level relative to team leaders in the U.S., as can be seen from Panel (B) of Figure 11. In the U.S., team leaders who just enter the inventor labor force produce on average patents with 2.1 forward citations in a 3-year window; in Germany, the average patent quality is almost one third less (0.7 forward citations). However, as panel (C) shows, the productivity of all inventors grows faster in Germany than in the U.S. Thus, not only do German inventors interact more often, they also learn more from each interaction. Finally, over the

[51] We use the same seed as in our the benchmark calibration.

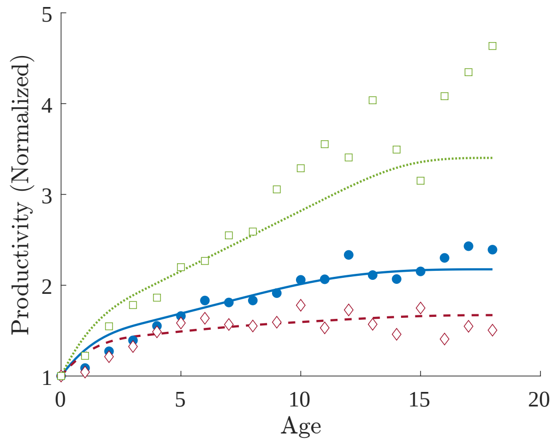
Figure 11: Targeted Moments for the U.S. and Germany



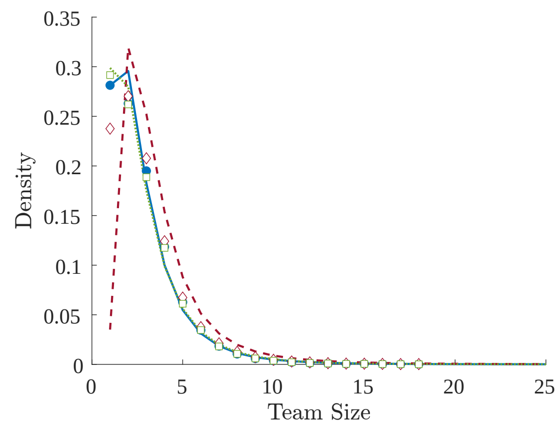
(a) High-Quality Interactions



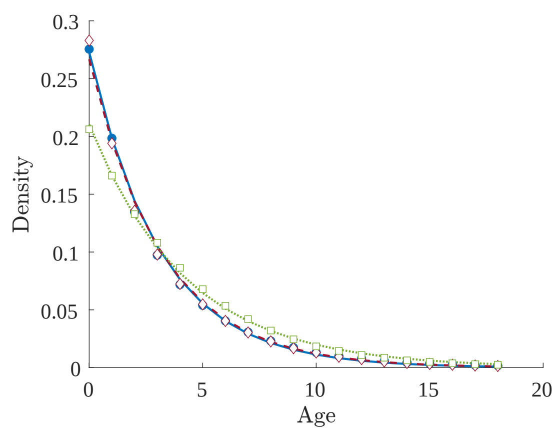
(b) Idea Quality q



(c) Productivity's Life Cycle for All Individuals



(d) Team Size Distribution



(e) Age Distribution

Panel (A) represents the high-quality interactions of a team leader as a function of his age (age 0 is the first year in the patent data). Panel (B): represents the de-trended idea quality $q(z,t)e^{-g_q t} = q(z,t)e^{-g(1-\eta)t}$, where $g_q = g(1-\eta)$ is the growth rate of idea quality, as a function of team leader's age. Panel (C) depicts average inventor productivity z across all inventors at a given age. It is normalized by productivity at age 0. Panel (D) shows the distribution of the team size n, and Panel (E) shows the distribution of age (time in the sample). For all panels, the blue solid lines represent the baseline case with the full sample, the maroon dashed lines the case using U.S. data and calibration, and the green dotted lines using DE data and calibration.

period 1980 to 2010, the U.S. has experienced a 2.74% annual growth in real GDP, which is higher than the 2.42% growth rate in Germany.[52]. This higher growth rate implies that the efficiency α of transforming innovations into output is higher for the US.

The data also shows that teams are on average larger in the U.S. (see Panel (D)). In line with this, there are fewer team leaders in the U.S. (50% of all inventors) than in Germany (54% of all inventors). In the U.S., superstar team leaders build larger teams – this more rigorous selection is reflected in the span of control η, which is higher in the U.S. (see Table 9).

The steeper lifecycle of productivity in Germany is entirely in line with the steeper profile of wages documented in Lagakos, Moll, Porzio, Qian, and Schoellman (2018b). These authors argue that the different life cycles are most likely explained by human capital and search frictions. For instance, manual occupations have flatter profiles than high-skilled occupations. Our model provides a micro-foundation for how human capital is built (through interactions) and an explanation of how these human capital acquisition processes differ in the U.S. and in Germany. Also consistent with Lagakos et al. (2018b), we found in Section 3 that interactions had a stronger effect on productivity in high-tech sectors.

Growth Decomposition US and Germany

First, we repeat the growth decomposition exercise for the US and Germany. As in the baseline, Table 12 panel A, shows the two learning channels are important for growth in both countries. In the US, external learning is as important as interactions; when either channel is shut down, growth decreases to 1.28%. In contrast, for Germany, interactions are the fundamental source for growth. If we shut down this channel growth positive growth is no longer possible, indicating that external sources alone do not suffice to overcome the death rate of inventors. This illustrates again the differences in the learning processes across the two countries. Additionally, the decomposition show the two learning channels are complementary in both countries and that allowing for endogenous interactions attenuates the effects of shutting down external learning, especially for the US.

Table 11: Growth Decomposition Countries and Model with Interactions for All

	Baseline	No interactions	No external learning
US (Fixed $m(x,s)$)	2.74%	1.28%	1.28%
US (Endogenous $m(x,s)$)	2.74%	1.28%	1.85%
Germany (Fixed $m(x,s)$)	2.42%	0%	1.77%
Germany (Endogenous $m(x,s)$)	2.42%	0%	2.00%

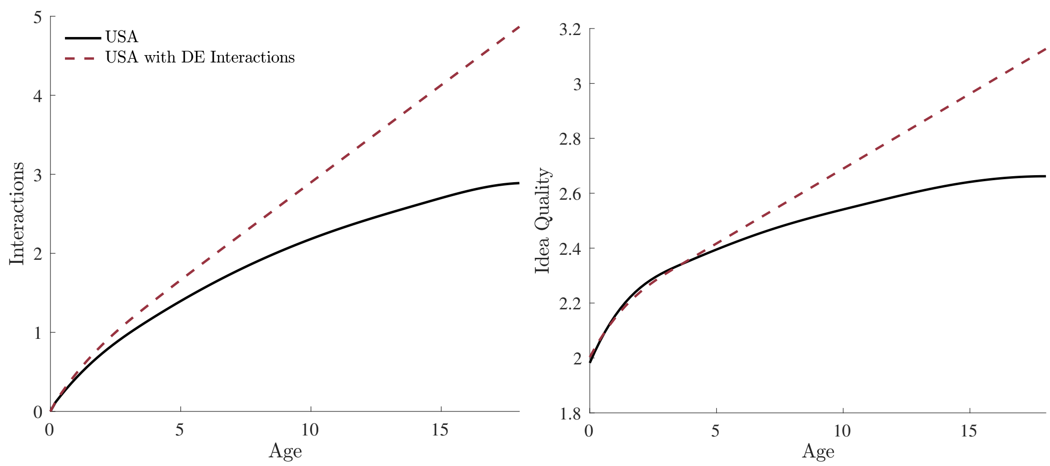
Note: This table estimates the growth rate when shutting down each of the learning channels using the calibration parameters in the US and Germany and for the model with interactions for all. For US and Germany has the effects while fixing the endogenous meeting rate to the one obtained in the baseline case and also after allowing agents to internalize the change in the parameters when optimizing $m(x,s)$.

[52]The real GDP data used to compute these annual growth rates comes from the expenditure-side real GDP at chained PPPs (in mil. 2011 US dollars), from the Penn World Table 9.0 available on `www.ggdc.net/pwt`.

Counterfactual analysis: Increasing interactions in the US to the German level

Suppose we reduce the cost of interactions $\kappa_{\text{US}}(s)$ to achieve Germany's interaction profile. To do so we fix parameters all parameter to the US and re-estimate the interaction cost function targeting Germany's interactions. The new estimated cost parameters $\kappa_0 = 0.92$, $\kappa_1 = -0.04$, $\kappa_{min} = 0.03$, imply that on average the interaction costs are reduced by 22%. This increases the optimal meeting rate by 35% on average, and the growth rate by 17%, from 2.74% to 3.21%. Figure 12 shows that the average number of interactions of team leaders increases at all ages as does average patent quality (see Appendix Figure A-6 for additional results). The life cycle profile of productivity in the baseline case steepens as human capital is built more extensively at each age through increased interactions.

Figure 12: Reducing Interaction Costs in the Benchmark to the Level in Germany



(a) Interactions (b) Average Patent Quality

Panel (A): represents the high-quality interactions of a team leader as a function of his age (age 0 is the first year in the patent data). Panel (B) depicts average patent quality $q(z, T)$. In each panel, the solid blue line represents the baseline case for the US. The black dashed line corresponds to the case changing the US interaction cost to target Germany's interactions.

5.5 Robustness: Interaction for All

In the model, team members learn from their team leader or more knowledgeable co-inventors. This is consistent with our measure of interactions in the data. However, our results are robust if we consider a model where both team members and team leaders draw from the full distribution (as in Akcigit et al. (2018)). Suppose all agents interact and learn from each other before forming teams, and that individuals only make a one time choice for the endogenous meeting rate m when they are age 0.[53]

[53] In our model we also generalize this choice allowing for the endogenous meeting rate $m(s)$ to adjust optimally as individual age and their learning opportunities change.

Table 12: Growth Decomposition Countries and Model with Interactions for All

	Baseline	No interactions	No external learning
Interactions for All (Fixed m)	2.50%	0.10%	0.62%

Note: This table estimates the growth rate when shutting down each of the learning channels for the model with interactions for all.

Table 12 shows the same main conclusions of the growth decomposition hold; both channels are important for growth, interactions are relatively more important, and there is strong complementarity between the two learning channels.[54] Quantitatively, if all individuals interact, external learning becomes more important, and the complementarity between the two learning channels is stronger. Better external learning opportunities further potentiate the effect of interactions when team leaders (who are high-productivity individuals) are also able to learn and interact with other agents endogenously.

6 Conclusion

A key input into the innovation production process is an inventor's own knowledge. That knowledge is built through interactions with others. In this paper, we build an innovation-based endogenous growth model, in which we micro-found the innovation process at the individual inventor and research team levels. An inventor acquires knowledge through interactions with others who are more knowledgeable than him or through alternative channels such as individual discovery, experience, or learning-by-doing. The most productive (knowledgeable) inventors form research teams of varying sizes and produce ideas of heterogeneous qualities that depend on their own knowledge. Thus, we bring together the diffusion and innovation-based growth literatures; in our paper, the knowledge diffusion model nests within the innovation-based growth model.

One of our key contributions is to bring data to a heretofore largely theoretical literature; we provide a framework that brings micro data to a general equilibrium innovation model. We use new panel data on millions of inventors matched to their employers and patents. This allows us to give empirical content to typically very hard to measure concepts such as interactions, productivity or knowledge, and research teams, to discipline the key ingredients of the model, and to estimate it. We find that interactions are crucial in explaining individual productivity and aggregate growth.

In future work, our model could be used to study the effects on growth and productivity of many important policies that affect the cost of interactions in different ways. What are the effects of non-compete laws which prevent inventors from easily moving between companies? Do labor market frictions and dynamics indirectly play a role for innovation and productivity because of their impact on the ease of interactions? How do immigration policies affect the inflow of new inventors and ideas? Our theoretical and quantitative framework, coupled with this new large-scale micro data could shed light on these questions.

[54]For the details of the interactions for all model and the results on other counterfactuals, see Akcigit et al. (2018).

References

Acemoglu, D., P. Aghion, and F. Zilibotti (2006). Distance to Frontier, Selection and Economic Growth. *Journal of the European Economic Association 4*(1), 37–74.

Aghion, P., U. Akcigit, and P. Howitt (2014). What do we learn from schumpeterian growth theory? In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*. Amsterdam: Elsevier.

Aghion, P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica 60*(2), 323–351.

Agrawal, A., A. Galasso, and A. Oettl (2017). Roads and innovation. *Review of Economics and Statistics 99*, 417–434.

Akcigit, U., S. Baslandze, and S. Stantcheva (2016). Taxation and the International Mobility of Inventors. *American Economic Review 106*(10), 2930–2981.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018). Dancing with the stars: Innovation through interactions. Technical report, National Bureau of Economic Research.

Akcigit, U. and N. Goldschlag (2025). Measuring the characteristics and employment dynamics of us inventors. *Journal of Economic Growth (Forthcoming)*.

Akcigit, U. and W. R. Kerr (2018). Growth through Heterogeneous Innovations. *Journal of Political Economy 126*(4), 1374–1443.

Alvarez, F., F. Buera, and R. Lucas (2008). Models of idea flows. *National Bureau of Economic Research Working Paper* (14135), 1–12.

Alvarez, F., F. Buera, and R. Lucas (2017). Idea flows, economic growth and trade. *University of Chicago Working Paper*.

Azoulay, P., C. Fons-Rosen, and J. G. Zivin (2019). Does science advance one funeral at a time? *American Economic Review 109*(8), 2889–2920.

Benhabib, J., J. Perla, and C. Tonetti (2014). Catch-up and fall-back through innovation and imitation. *Journal of Economic Growth 19*, 1–35.

Benhabib, J., J. Perla, and C. Tonetti (2021). Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier. *Econometrica 89*(5), 2261–2301.

Bloom, N., R. Sadun, and J. V. Reenen (2012). Americans do i.t. better: Us multinationals and the productivity miracle. *American Economic Review 102*(1), 167–201.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4), 1347–1393.

Breschi, S., F. Lissoni, and G. Tarasconi (2016). Inventor data for research on migration and innovation: The ethnic-inv pilot database. In C. Fink and E. Miguelez (Eds.), *The International Mobility of Talent and Innovation: New Evidence and Policy Implications*. Cambridge Univ. Press.

Buera, F. and E. Oberfield (2020). The global diffusion of ideas. *Econometrica 88*(1), 83–114.

Buera, F. J. and J. R. E. Lucas (2018). Idea flows and economic growth. *Annual Review of Economics 10*, 315–345.

Caicedo, S., R. E. Lucas Jr, and E. Rossi-Hansberg (2019). Learning, career paths, and the distribution of wages. *American Economic Journal: Macroeconomics 11*(1), 49–88.

Coffano, M. and G. Tarasconi (2014). Crios-patstat database: Sources, contents and access rules. *Center for Research on Innovation, Organization and Strategy, CRIOS Working Paper* (1).

De Rassenfosse, G., W. E. Griffiths, A. B. Jaffe, and E. Webster (2021). Low-quality patents in the eye of the beholder: Evidence from multiple examiners. *The Journal of Law, Economics, and Organization 37*(3), 607–636.

Grossman, G. and E. Helpman (1991). Quality ladders in the theory of growth. *The Review of Economic Studies 58*(1), 43–61.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The nber patent citation data file: Lessons, insights and methodological tools. National Bureau of Economic Research WP 8498.

Herkenhoff, K., J. Lise, G. Menzio, and G. M. Phillips (2024). Production and learning in teams. *Econometrica 92*(2), 467–504.

Hoisl, K. (2007). Tracing mobile inventors: The causality between inventor mobility and inventor productivity. *Research Policy 36*(5), 619–636.

Jaravel, X., N. Petkova, and A. Bell (2018). Team-specific capital and innovation. *American Economic Review 108*(4-5), 1034–73.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021). Learning from coworkers. *Econometrica 89*(2), 647–676.

Jones, B. F. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *Review of Economic Studies 76*(1).

Jones, B. F. (2010). Age and great invention. *Review of Economics and Statistics 92*(1), 1–14.

Jones, B. F. and B. Weinberg (2011). Age Dynamics in Scientific Creativity. *Proceedings of the National Academy of Sciences 108*(47), 18855–19096.

Jones, B. F., S. Wuchty, and B. Uzzi (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science 322*(5905), 1259–1262.

Jones, C. I. (1995). R&d-based models of economic growth. *Journal of Political Economy*, 759–784.

Jovanovic, B. and R. Rob (1989). The growth and diffusion of knowledge. *The Review of Economic Studies 56*(4), 569–582.

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy 112*(5), 986–1018.

König, M. D., J. Lorenz, and F. Zilibotti (2016). Innovation vs. imitation and the evolution of productivity distributions. *Theoretical Economics 11*(3), 1053–1102.

Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica 65*(6), 1389–1419.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018a). Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants. *Journal of Human Capital, forthcoming 12*(2), 305–342.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018b). Life-Cycle wage Growth Across Countries. *Journal of Political Economy 126*(2), 797–849.

Lemieux, T. (2006). The mincer equation thirty years after schooling, experience, and earnings. In S. Grossbard-Schechtman (Ed.), *Jacob Mincer, A Pioneer of Modern Labor Economics*, pp. 127–145. Springer.

Li, G.-C., R. Lai, A. D'Amour, D. M. Doolin, Y. Sun, V. I. Torvik, Z. Y. Amy, and L. Fleming (2014). Disambiguation and co-authorship networks of the us patent inventor database (1975–2010). *Research Policy 43*(6), 941–955.

Lissoni, F., M. Coffano, A. Maurino, M. Pezzoni, and G. Tarasconi (2010). Ape-inv's "name game" algorithm challenge: A guideline for benchmark data analysis & reporting. *Version 1*, 875–889.

Lucas, R. E. and B. Moll (2014). Knowledge growth and the allocation of time. *Journal of Political Economy 122*(1), 1–51.

Lucas, R. E. J. (1978). On the size distribution of business firms. *Bell Journal of Economics*, 508–523.

Lucas, R. J. (2009). Ideas and growth. *Economica 76*, 1–19.

Luttmer, E. G. (2007). Selection, Growth, and the Size Distribution of Firms. *Quarterly Journal of Economics 122*(3), 1103–1144.

Luttmer, E. G. J. (2012). Eventually, noise and imitation implies balanced growth. *Federal Reserve Bank of Minneapolis Working Paper* (699).

Luttmer, E. G. J. (2015). An assigment model of knowledge diffusion and income inequality. *Federal Reserve Bank of Minneapolis Working Paper* (509).

Miguelez, E. (2018). Inventor diasporas and the internationalization of technology. *The World Bank Economic Review 32*(1), 41–63.

Miguelez, E., C. Fink, et al. (2013). *Measuring the international mobility of inventors: A new database*, Volume 8. WIPO.

Nakajima, R., R. Tamura, and N. Hanaki (2010). The effect of collaboration network on inventors' job match, productivity and tenure. *Labour Economics 17*(4), 723–734.

Pakes, A. S. (1986). Patents as options: Some estimates of the value of holding european patent stocks. *Econometrica 54*(4), 755–84.

Perla, J. and C. Tonetti (2014). Equilibrium imitation and growth. *Journal of Political Economy 122*(1), 52–76.

Perla, J., C. Tonetti, and M. E. Waugh (2021). Equilibrium technology diffusion, trade, and growth. *American Economic Review 111*(1), 73–128.

Pezzoni, M., F. Lissoni, and G. Tarasconi (2014). How to kill inventors: testing the massacrator© algorithm for inventor disambiguation. *Scientometrics 101*, 477–504.

Raffo, J. and S. Lhuillery (2009). How to play the "names game": Patent retrieval comparing different heuristics. *Research Policy 38*(10), 1617–1627.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy 98*(5), 71–102.

Schankerman, M. and A. Pakes (1986). Estimates of the value of patent rights in european countries during the post-1950 period. *Economic Journal 96*(384), 1052–1076.

Staley, M. (2011). Growth and the diffusion of ideas. *Journal of Mathematical Economics 47*, 470–478.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *Rand Journal of Economics*, 172–187.

Wuchty, S., B. F. Jones, and B. Uzzi (2007). The increasing dominance of teams in the production of knowledge. *Science 316*(5827), 1036–1039.

Appendix

A-1 Proofs from Section 2

A-1.1 Value function of final good producers

We here show that the value function of any final good producer is linear in A_i. With some probability χ (which we do not need to specify here), the final good producer will purchase an idea of quality q for a total price $P = pq$ from the research teams. The value of a final good producer satisfies:

$$rV_i(t) - \dot{V}_i(t) \quad = \quad \Pi_i(t) + \chi \mathbb{E}(V_i(A_i + q) - V_i(A_i) - P) \tag{A-1}$$

where the expectation is over the possible values of q. Assume that on a balanced growth path, the value V_i grows at rate g_V so that $\dot{V}_i(t) = g_V V_i(t)$. The wage $w_u(t)$ grows at rate g_w and $\bar{A}$ grows at rate $g_{\bar{A}}$.

Profits are: $\Pi_i(t) = \alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} w_u(t)^{\frac{\alpha-1}{\alpha}} A_i(t)$.

Let us conjecture that the value function takes the form:

$$V_i(A_i) = \hat{v} A_i w_u^{\frac{\alpha-1}{\alpha}} + \tilde{v} \bar{A}^{\alpha} \tag{A-2}$$

for some coefficients $\hat{v}$ and $\tilde{v}$.

Under this conjecture, the left hand side of A-1 becomes:

$$r\hat{v} A_i w_u^{\frac{\alpha-1}{\alpha}} + r\tilde{v}\bar{A}^{\alpha} - \left(\frac{\alpha-1}{\alpha} g_w\right) \hat{v} A_i w_u^{\frac{\alpha-1}{\alpha}} - \alpha g_{\bar{A}} \tilde{v} \bar{A}^{\alpha}$$

Under the assumptions in the main text, the total price for an idea of quality q is $P = \beta \frac{\hat{v} q}{\bar{A}(t)^{1-\alpha}}$, or, using the formula for the wage, $P = \tilde{\Omega} q w_u^{\frac{\alpha-1}{\alpha}}$ where $\tilde{\Omega}$ is a constant. The right hand side of (A-1) is then:

$$\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}} w_u(t)^{\frac{\alpha-1}{\alpha}} A_i(t) + \chi \mathbb{E}_q \left[\hat{v} q w_u^{\frac{\alpha-1}{\alpha}} - \tilde{\Omega} q w_u^{\frac{\alpha-1}{\alpha}} \right]$$

Equating the coefficients on all terms in $A_i(t)$ on the left hand side and right hand side we get:

$$\hat{v} = \frac{\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{r - \left(\frac{\alpha-1}{\alpha} g_w\right)}$$

Note that:

$$x\mathbb{E}_q \left[\hat{v} q w_u^{\frac{\alpha-1}{\alpha}} - \tilde{\Omega} q w_u^{\frac{\alpha-1}{\alpha}} \right] = \chi \left[\hat{v} - \tilde{\Omega}\right] \bar{q} w_u^{\frac{\alpha-1}{\alpha}}$$

Note that from the formula for the wage: $\frac{1}{1-\alpha} w_u(t) = \bar{A}(t)^{\alpha}$. In addition, $\bar{q} = g_{\bar{A}} \bar{A}$.

Equating all remaining terms and substituting for $\bar{A}$ yields:

$$\begin{aligned} r\tilde{v}\frac{1}{1-\alpha}w_u - \alpha g_{\bar{A}}\tilde{v}\frac{1}{1-\alpha}w_u &= \chi\left[\hat{v}-\tilde{\Omega}\right]g_{\bar{A}}\left(\frac{1}{1-\alpha}\right)^{\frac{1}{\alpha}} w_u^{\frac{1}{\alpha}} w_u^{\frac{\alpha-1}{\alpha}} \\ \tilde{v} &= \frac{\chi\left[\hat{v}-\tilde{\Omega}\right]g_{\bar{A}}}{(r-\alpha g_{\bar{A}})(1-\alpha)^{\frac{1-\alpha}{\alpha}}} \end{aligned}$$

The value function thus takes the shape conjectured in (A-2) and is an affine function of $A_i(t)$.

We can further rewrite the value function as a function of own productivity $A_i(t)$ and aggregate productivity $\bar{A}(t)$ only. Using that $w_u(t)^{\frac{\alpha-1}{\alpha}} = \frac{1}{\bar{A}^{1-\alpha}}(1-\alpha)^{\frac{\alpha-1}{\alpha}}$, we get that: $V_i(A_i) = \frac{\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}}}{r-\left(\frac{\alpha-1}{\alpha}g_w\right)} A_i \frac{1}{\bar{A}^{1-\alpha}}(1-\alpha)^{\frac{\alpha-1}{\alpha}} + \tilde{v}\bar{A}^{\alpha}$, which simplifies to:

$$V_i(A_i) = \frac{\alpha}{r+\left(\frac{1-\alpha}{\alpha}g_w\right)}\frac{A_i}{\bar{A}^{1-\alpha}} + \tilde{v}\bar{A}^{\alpha} = v\frac{A_i}{\bar{A}^{1-\alpha}} + \tilde{v}\bar{A}^{\alpha} \qquad \text{with:} \qquad v = \frac{\alpha}{r+\frac{1-\alpha}{\alpha}g_w}$$

This means that the marginal return to purchasing more quality ($\frac{dV_i(A_i)}{dA_i}$) is the same across all final good producers, regardless of their starting level of A_i. This confirms the claim in the main text.

A-1.2 Balance Growth Path Computations

In this section we describe the detailed BGP computations and the proof of Proposition 1.

First, we can write down the HJB equations (3) and (5) on a BGP,

$$\begin{aligned} g v_l'(x,s)x &= \pi(x) + m_X\int_x^{\infty}\left(\iota v_l(\tilde{x},s)+(1-\iota)v_m(\tilde{x},s)-v_l(x,s)\right)\psi(\tilde{x},s)d\tilde{x} \\ &\quad + (1-\iota)(v_m(x,s)-v_l(x,s)) - (\rho+\delta-g(1-\eta))v_l(x,s) \qquad \text{(A-3)} \\ g v_m'(x,s)x &= w_0 \\ &\quad + \max_m\left\{m\int_x^{\infty}\left(\iota v_l(\tilde{x},s)+(1-\iota)v_m(\tilde{x},s)-v_m(x,s)\right)\phi(\tilde{x})d\tilde{x} - c(m,s)\right\} \\ &\quad + m_X\int_x^{\infty}\left(\iota v_l(\tilde{x},s)+(1-\iota)v_m(\tilde{x},s)-v_m(x,s)\right)\psi(\tilde{x},s)d\tilde{x} \\ &\quad + \iota(v_l(x,s)-v_m(x,s)) - (\rho+\delta-g(1-\eta))v_m(x,s). \qquad \text{(A-4)} \end{aligned}$$

The indifference condition and the market clearing condition that determine the productivity cutoffs $\hat{x}(s)$ and the wage w_0 are given by:

$$v_l(\hat{x}(s),s) = v_m(\hat{x}(s),s)\,, \qquad \text{(A-5)}$$

$$\int_0^{\infty}\left[\Phi(\hat{x}(s)) + (1-\iota)(1-\Phi(\hat{x}(s)))\right]d\Omega(s) = \iota\int_0^{\infty}\int_{\hat{x}(s)}^{\infty}\left(\frac{p_0\eta}{w_0}\right)^{1/(1-\eta)} x\phi(x)dx\, d\Omega(s). \qquad \text{(A-6)}$$

From equation (A-4) and using the interaction costs from Assumption 2, $c(m,s) = \frac{\kappa(s)}{2}m^2$, we solve for the optimal meeting rate,

$$m^*(x,s) = \frac{1}{\kappa(s)}\int_x^{\infty}\left(\iota v_l(\tilde{x},s)+(1-\iota)v_m(\tilde{x},s)-v_m(x,s)\right)\phi(\tilde{x})d\tilde{x}$$

If $x \leq \hat{x}(s)$ individuals choose to be team members even if they have a research idea. This implies the gains from endogenous interactions come from meeting individuals with productivity larger than the cutoffs. So for team members below the cutoff, $x \leq \hat{x}(s)$,

$$m^*(s) = \frac{1}{\kappa(s)} \int_{\hat{x}(s)}^{\infty} \left(\iota v_l(\tilde{x}, s) + (1-\iota) v_m(\tilde{x}, s) - v_m(s) \right) \phi(\tilde{x}) d\tilde{x} \tag{A-7}$$

Now turning to the dynamics of the productivity distribution, the KFE equation for individuals of a cohort born in period t_b, of age s at time t with productivity below the cutoff $\hat{z}(s,t)$ is,

$$\begin{aligned}\frac{\partial \tilde{h}(z,s,t_b)}{\partial s} = & \underbrace{-m(z,s)(1-F(z,t))\tilde{h}(z,s,t_b) + m(z,s)\tilde{H}(z,s,t_b)f(z,t)}_{\text{Learning from team leaders + co-inventors}} \\ & \underbrace{-m_X(1-E(z,s,t))\tilde{h}(z,s,t_b) + m_X\tilde{H}(z,s,t_b)e(z,s,t)}_{\text{External learning}}, \quad \forall z < \hat{z}(s,t), \ \forall s\end{aligned}$$

For inventors of age s at time t, born in period t_b above the cutoffs, $z(t) \geq \hat{z}(s,t)$, it evolves as,

$$\begin{aligned}\frac{\partial \tilde{h}(z,s,t_b)}{\partial s} = & \underbrace{\tilde{H}(\hat{z},s,t_b)(m(z,s)f(z,t) + m_X e(z,s,t))}_{\text{Inflow of team members with } z<\hat{z}} \\ & \underbrace{+m(z,s)(1-\iota)\left(\tilde{H}(z,s,t_b) - \tilde{H}(\hat{z},s,t_b)\right) f(z,t)}_{\text{Inflow of team members with } z \geq \hat{z}} \underbrace{-m(z,s)(1-\iota)\left(1-F(z,t)\right)\tilde{h}(z,s,t_b)}_{\text{Outflow of team members with } z \geq \hat{z}} \\ & \underbrace{-m_X\left(1-E(z,s,t)\right)\tilde{h}(z,s,t_b) + m_X\left(\tilde{H}(z,s,t_b) - \tilde{H}(\hat{z}(s,t),s,t_b)\right)e(z,s,t)}_{\text{External learning}}, \quad \forall z \geq \hat{z}(s,t), \ \forall s.\end{aligned}$$

If we integrate the KFE equations between age $[0,s]$, we get,

$$\begin{aligned}\tilde{h}(z,s,t_b) = & -\int_0^s m(z,\tau)\left(1-F(z,t_b+\tau)\right)\tilde{h}(z,\tau,t_b)d\tau + \int_0^s m(z,\tau)\tilde{H}(z,\tau,t_b)f(z,t_b+\tau)d\tau \\ & - m_X\int_0^s \left(1-E(z,\tau,t_b+\tau)\right)\tilde{h}(z,\tau,t_b)d\tau + m_X\int_0^s \tilde{H}(z,\tau,t_b)e(z,\tau,t_b+\tau)d\tau \\ & + \tilde{h}(z,0,t_b), \quad \forall z < \hat{z}(s,t), \ \forall s.\end{aligned} \tag{A-8}$$

And for inventors above the cutoffs, $z(t) \geq \hat{z}(s,t)$,

$$\begin{aligned}\tilde{h}(z,s,t_b) = & \int_0^s m(z,\tau)\tilde{H}(\hat{z}(\tau),\tau,t_b)f(z,t_b+\tau)d\tau - (1-\iota)\int_0^s m(z,\tau)\left(1-F(z,t)\right)\tilde{h}(z,s,t_b)d\tau \\ & + (1-\iota)\int_0^s m(z,\tau)\left(\tilde{H}(z,\tau,t_b) - \tilde{H}(\hat{z}(\tau),\tau,t_b)\right)f(z,t_b+\tau)d\tau \\ & - m_X\int_0^s \left(1-E(z,\tau,t_b+\tau)\right)\tilde{h}(z,\tau,t_b)d\tau + m_X\int_0^s \tilde{H}(z,\tau,t_b)e(z,\tau,t_b+\tau)d\tau \\ & + \tilde{h}(z,0,t_b), \quad \forall z \geq \hat{z}(s), \ \forall s.\end{aligned} \tag{A-9}$$

We can use the change of variable $t_b = t - s$ to map age and birth data into calendar time. Let $H(z,s,t) := \tilde{H}(z,s,t_b)$ denote the cumulative probability distribution of an individual of age s at calendar date t (with density $h(z,s,t) := \tilde{h}(z,s,t_b)$). This is simply a convenient change in notation, we are still following the dynamics of the productivity of different a cohort as it ages.

On a BGP recall the de-trended productivity is defined as $x = ze^{-gt}$, so,

$$\Gamma(x,s) = H(z,s,t) \qquad \text{and} \qquad \Phi(x) = F(z,t).$$

Taking the derivative with respect to x, we get the densities of these distributions on a BGP,

$$\zeta(x,s) = h(z,s,t)e^{gt} \ , \ \ \phi(x) = f(z,t)e^{gt}.$$

We also assume that on a BGP the external distribution is invariant so,

$$\Psi(x,s) = E(z,s,t) \ , \ \ \psi(x,s) = e(z,s,t)e^{gt}.$$

Note using these definitions,

$$(1 - F(z, t-s+\tau)) = 1 - \Phi\left(xe^{g(s-\tau)}\right) \quad , \quad f(z,t-s+\tau) = \phi(xe^{g(s-\tau)})e^{-g(t-s+\tau)}$$
$$(1 - E(z,\tau, t-s+\tau)) = 1 - \Psi\left(xe^{g(s-\tau)},\tau\right) \quad , \quad e(z,\tau,t-s+\tau) = \psi(xe^{g(s-\tau)},\tau)e^{-g(t-s+\tau)}$$

We can use replace these derivations in equations (A-8) and (A-9), to get the KFE on a BGP,

$$\zeta(x,s) = -\int_0^s m(x,\tau)\left(1-\Phi\left(xe^{g(s-\tau)}\right)\right)\zeta(x,\tau)d\tau + \int_0^s m(x,\tau)\Gamma(x,\tau)\phi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau$$
$$- m_X\int_0^s \left(1-\Psi\left(xe^{g(s-\tau)},\tau\right)\right)\psi(xe^{g(s-\tau)})e^{g(s-\tau)} + \zeta(xe^{gs},0)e^{gs} \ , \ \ \forall x < \hat{x}(s), \ \forall s. \qquad \text{(A-10)}$$

Similarly for $x \geq \hat{x}(s)$ we get,

$$\zeta(x,s) = \int_0^s m(x,\tau)\Gamma(\hat{x}(\tau),\tau)\phi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau - (1-\iota)\int_0^s m(x,\tau)\left(1-\Phi\left(xe^{g(s-\tau)}\right)\right)\zeta(x,\tau)d\tau$$
$$+ (1-\iota)\int_0^s m(x,\tau)\left(\Gamma(x,\tau) - \Gamma(\hat{x}(\tau),\tau)\right)\phi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau$$
$$- m_X\int_0^s\left(1-\Psi\left(xe^{g(s-\tau)},\tau\right)\right)\zeta(x,\tau)d\tau + m_X\int_0^s \Gamma(z,\tau)\psi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau$$
$$+ \zeta(xe^{gs},0)e^{gs} \ , \ \ \forall x \geq \hat{x}(s), \ \forall s. \qquad \text{(A-11)}$$

The de-trended aggregate distributions are,

$$\Phi(x) = \int_0^\infty \Gamma(x,s)d\Omega(s) \ \ , \ \ \phi(x) = \int_0^\infty \zeta(x,s)d\Omega(s). \qquad \text{(A-12)}$$

To derive the relation between the growth rate of the economy and the meeting rates $m(s)$ and m_X we make the additional assumption that on a BGP, the external and the age zero distributions have a common Pareto tail parameters, θ, i.e.

$$\lim_{x\to\infty}\frac{1-\Psi(x,s)}{x^{-1/\theta}} = \varrho(s) \ , \ \lim_{x\to\infty}\frac{1-\Gamma(x,0)}{x^{-1/\theta}} = k_0$$

for some $\theta, k_0 > 0$, and $\varrho(s) > 0 \ \forall s$.
Finally, suppose $\lim_{z\to\infty}\frac{1-F(z,0)}{z^{-1/\theta}} = \lim_{z\to\infty}\frac{1-\Phi(x)}{x^{-1/\theta}} = k < \infty$, we can show the cross-sectional productivity distribution $\Phi(x)$ has a Pareto tail θ, so $k > 0$, and derive an expression for the growth rate g.

A-1.3 Proof Proposition 1

Proof. Let $k = \lim_{x\to\infty} \frac{1-\Phi(x)}{x^{-1/\theta}}$. Suppose that the initial distribution has a finite Pareto tail limit, so on a BGP, $\lim_{z\to\infty} \frac{1-F(z,0)}{z^{-1/\theta}} = \lim_{x\to\infty} \frac{1-\Phi(x)}{x^{-1/\theta}} < \infty$. We will show that this limit is positive, i.e. $k > 0$.

Since we are taking the limit as $x \to \infty$ we consider the KFE for team leaders. Combining equations (A-11) and (A-12), we get:

$$\begin{aligned}
\phi(x) = &\int_0^\infty \int_0^s m(x,\tau)\Gamma(\hat{x}(\tau),\tau)\phi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau d\Omega(s) \\
&- (1-\iota)\int_0^\infty \left[\int_0^s m(x,\tau)\left(1-\Phi\left(xe^{g(s-\tau)}\right)\right)\zeta(x,\tau)d\tau - \int_0^s m(x,\tau)\left(\Gamma(x,\tau)-\Gamma(\hat{x}(\tau),\tau)\right)\phi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau\right]d\Omega(s) \\
&- m_X\int_0^\infty \left[\int_0^s \left(1-\Psi\left(xe^{g(s-\tau)},\tau\right)\right)\zeta(x,\tau)d\tau + m_X\int_0^s \Gamma(z,\tau)\psi(xe^{g(s-\tau)})e^{g(s-\tau)}d\tau\right]d\Omega(s) \\
&+ \int_0^\infty \zeta(xe^{gs},0)e^{gs}d\Omega(s).
\end{aligned} \tag{A-13}$$

Now, using Assumption 3 on the Pareto tail of the distributions and l'Hospital rule, $\lim_{x\to\infty} \frac{\phi(x)}{x^{-1/\theta-1}} = \frac{k}{\theta}$ and $\lim_{x\to\infty} \frac{\zeta(x,\tau)}{x^{-1/\theta-1}} = \frac{k(\tau)}{\theta}$. It follows,

$$\begin{aligned}
&\lim_{x\to\infty} \frac{1-\Phi\left(xe^{g(s-\tau)}\right)}{x^{-1/\theta}} = ke^{-\frac{g}{\theta}(s-\tau)}, \quad \lim_{x\to\infty} \frac{1-\Psi\left(xe^{g(s-\tau)},\tau\right)}{x^{-1/\theta}} = \varrho(\tau)e^{-\frac{g}{\theta}(s-\tau)}, \\
&\lim_{x\to\infty} \frac{\zeta(xe^{gs},0)}{x^{-1/\theta-1}} = \frac{k_0e^{-\frac{g}{\theta}s-gs}}{\theta}, \quad \lim_{x\to\infty} \frac{\phi(xe^{g(s-\tau)})}{x^{-1/\theta-1}} = \frac{ke^{-\frac{g}{\theta}(s-\tau)-g(s-\tau)}}{\theta}, \\
&\lim_{x\to\infty} \frac{\psi(xe^{g(s-\tau)},\tau)}{x^{-1/\theta-1}} = \frac{\varrho(\tau)e^{-\frac{g}{\theta}(s-\tau)-g(s-\tau)}}{\theta}.
\end{aligned}$$

Using these properties for large x, equation (A-13) implies,

$$\begin{aligned}
\frac{k}{\theta}x^{-1/\theta-1} = &\int_0^\infty \int_0^s m(\tau)\Gamma(\hat{x}(\tau),\tau)\frac{ke^{-\frac{g}{\theta}(s-\tau)}}{\theta}x^{-1/\theta-1}d\tau d\Omega(s) \\
&- (1-\iota)\int_0^\infty \int_0^s m(\tau)\frac{ke^{-\frac{g}{\theta}(s-\tau)}x^{-1/\theta}}{\left(1-\Phi\left(\hat{x}e^{g(s-\tau)}\right)\right)^2}\frac{k(\tau)}{\theta}x^{-1/\theta-1}d\tau d\Omega(s) \\
&+ (1-\iota)\int_0^\infty \int_0^s m(\tau)\left(1-\Gamma(\hat{x}(\tau),\tau)\right)\frac{ke^{-\frac{g}{\theta}(s-\tau)}}{\theta}x^{-1/\theta-1}d\tau d\Omega(s) \\
&- m_X\int_0^\infty \int_0^s \varrho(\tau)e^{-\frac{g}{\theta}(s-\tau)}x^{-1/\theta}\frac{k(\tau)}{\theta}x^{-1/\theta-1}d\tau d\Omega(s) \\
&+ m_X\int_0^\infty \int_0^s \frac{\varrho(\tau)e^{-\frac{g}{\theta}(s-\tau)}}{\theta}x^{-1/\theta-1}d\tau d\Omega(s) \\
&+ \int_0^\infty \zeta(xe^{gs},0)\frac{k_0e^{-\frac{g}{\theta}s}}{\theta}x^{-1/\theta-1}d\Omega(s).
\end{aligned}$$

Moreover, from (A-7), for $x \leq \hat{x}(s)$, $m(x,\tau) = m^*(\tau)$ and $\lim_{x\to\infty} m^*(x,\tau) = 0\ \forall\tau$. So dividing both sides by $\frac{k}{\theta}x^{-1/\theta-1}$ and taking the limit $x \to \infty$,

$$k = k \int_0^\infty e^{-\frac{g}{\theta}s} \int_0^s m^*(\tau)\Gamma(\hat{x}(\tau),\tau) e^{\frac{g}{\theta}\tau} d\tau d\Omega(s)$$
$$+ m_X \int_0^\infty e^{-\frac{g}{\theta}s} \int_0^s \varrho(\tau) e^{\frac{g}{\theta}\tau} d\tau d\Omega(s)$$
$$+ k_0 \int_0^\infty e^{-\frac{g}{\theta}s} d\Omega(s) \tag{A-14}$$

From Assumption 3, $k_0 > 0$ and $\varrho(\tau) > 0$, $\forall\tau$. Given $m_X \geq 0$, equation (A-14) implies $k > 0$. Finally, dividing by k both sides of equation (A-14) we get the expression in Proposition 1. □

A-2 Jacobian Matrix

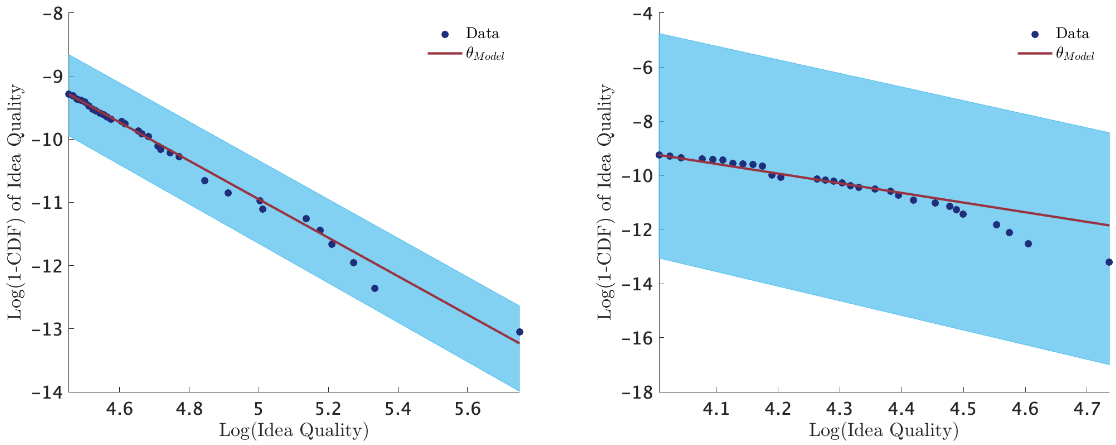
Table A-1: Marginal Effect of Parameters on Endogenous Meeting and Moments

	m_X	λ	θ	ν	η	δ	$\frac{k_0}{k}$	ι	κ_0	κ_1	$\kappa_{\min}$
Endogenous Meetings											
Average Meeting Rate, m^*	-0.21	0.24	0.26	-0.11	-0.13	-0.27	-0.11	-0.18	**-1.40**	0.21	-0.12
Average Estimated Moments											
Age Distribution	0.00	0.00	0.00	0.00	0.00	**-1.99**	0.00	0.00	0.00	0.00	0.00
Team Size	0.06	-0.09	**1.88**	0.05	-0.06	-0.20	-0.07	-0.43	0.05	-0.08	-0.08
Interactions	-0.03	0.17	**1.22**	-0.25	0.29	**-0.63**	-0.20	-0.15	**-1.00**	-0.11	-0.18
Idea Quality	0.08	0.20	**0.62**	0.12	0.06	0.20	-0.04	0.11	0.18	-0.03	-0.01
Productivity (Normalized)	**0.58**	-0.03	0.37	0.17	0.49	0.30	-0.23	0.41	**1.05**	-0.29	0.02
Fraction of Team Leaders	0.11	0.09	**-0.56**	-0.01	0.02	0.01	0.07	**0.67**	-0.03	0.01	0.02
Interactions Coefficient, β_1	-0.16	0.33	**-2.13**	**0.66**	0.00	**1.80**	-0.16	**-1.80**	**1.31**	0.00	0.33
Age Coefficient, β_2	**1.02**	0.26	**-1.02**	0.00	-0.26	**1.28**	0.26	**-1.53**	**1.02**	**0.51**	0.26

Note: This table contains the elasticity of the average endogenous meeting rate and the average value of the targeted moments with respect to each of the parameters around the estimated value. We compute the percentage change in the average of each moment after a 1% increase in the value of each parameter.

A-3 Additional Figures from Section 4.4

Figure A-1: Tail of idea Quality Distribution

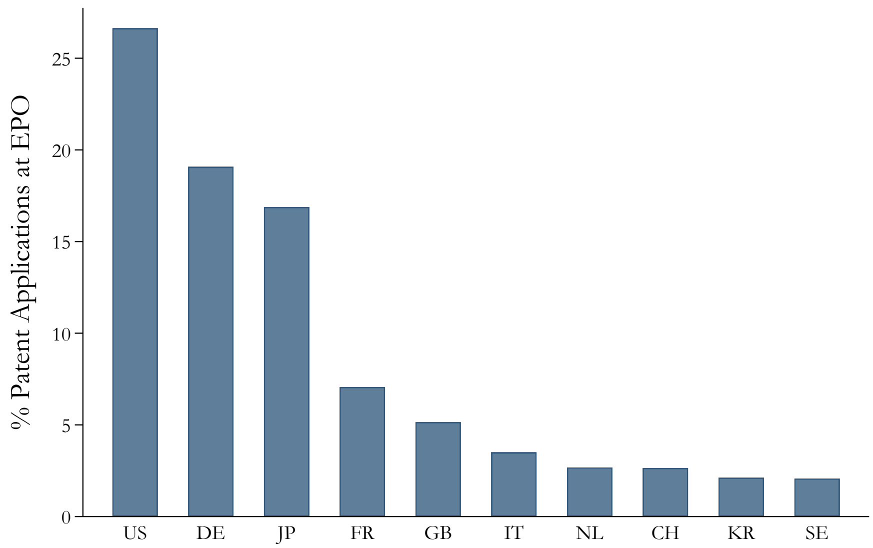


(a) United States Bootstrap CI (b) Germany Bootstrap CI

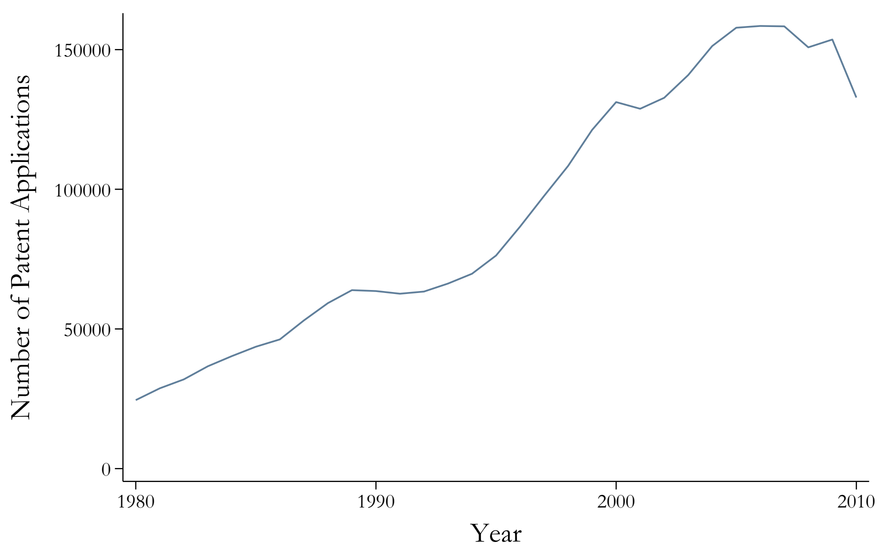
Notes: Panel (A) shows the tail of idea quality for the US. Panel (B) shows the tail of idea quality for Germany. Both figures plot the log of the counter CDF (log(1-CDF)) of the log idea quality $log(q)$ for idea quality above the 99.9th percentile. The red solid line is linear relation predicted by the model estimated θ anchoring the initial point to the on in the data. The shaded area are 95% confidence intervals of a regression using data between the 99.9th and the 99.9999th percentile.

A-4 Additional Data Information

Figure A-2: Applications at the EPO by Country and Year



(a) Applications across Countries



(b) Applications by Year

Notes: Panel (A) shows the percentage of all applications at the EPO by country. Panel (B) shows total applications at the EPO by year.

A-5 Additional Tables from Section 3.3

Table A-2: Results using USPTO Data

	Benchmark		High Tech Sector	Broader Interactions Measures	
				Firm	Region
High Quality Interactions (t-1)	0.016***	0.017***	0.019***	0.000020***	0.0000044***
	(0.00053)	(0.00056)	(0.00071)	(0.00000042)	(0.000000080)
Low Quality Interactions (t-1)		-0.00024***			
		(0.000069)			
Team Size	0.0085***	0.0085***	0.013***	0.0085***	0.0092***
	(0.00072)	(0.00072)	(0.0010)	(0.00072)	(0.00072)
Log Firm Size	-0.018***	-0.018***	-0.026***		-0.014***
	(0.0034)	(0.0034)	(0.0054)		(0.0034)
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year x Sector FE	Yes	Yes	Yes	Yes	Yes
Year x Region FE	Yes	Yes	Yes	Yes	Yes
Sector x Region FE	Yes	Yes	Yes	Yes	Yes
Team Leader FE	Yes	Yes	Yes	Yes	Yes
N	2993870	2993870	1474943	2950612	2992144
adj. R^2	0.19	0.19	0.20	0.19	0.19
F	214.7	175.8	184.0	627.4	662.6

Notes: All models include, in addition, sector, region, and year fixed effects (not listed). Column (1) uses the benchmark measure of interactions (past co-inventors). Column (2) also includes lower-quality past co-inventors. Column (3) limits the sample to the high-tech sector. Columns (4) and (5) consider, respectively, interactions measured as inventors in the same firm or in the same region. Standard errors clustered at the team leader are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

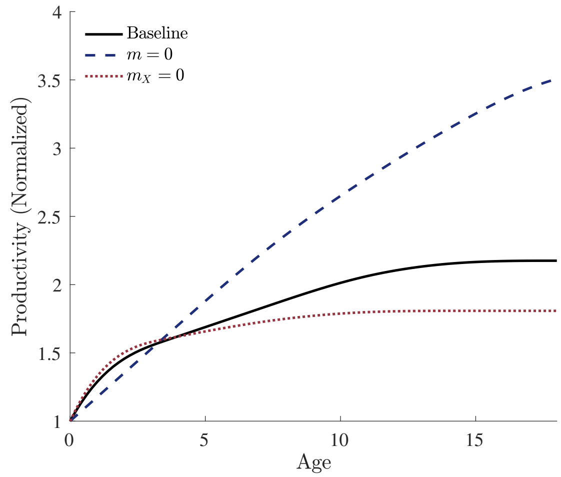
Table A-3: Summary Statistics–Switcher Team Leaders

	Mean	Standard Deviation	p50	p75	p90	p95	min	max
Before Becoming a Team Leader								
High-Quality Interactions	1.5	2.6	1	2	4	6	0	51
Low-Quality Interactions	3.0	3.8	2	4	7	9	0	95
Firm Size	535	898	131	651	1,610	2,462	1	7,634
Region Size	4,813	9,067	926	3,945	15,419	29,653	1	40,504
Distinct Firms	0.121	0.366	0	0	1	1	0	8
Distinct Regions	0.136	0.496	0	0	0	1	0	15
After Becoming a Team Leader								
Switches	1.4	0.8	1	1	2	3	1	15
Δ High Quality Interactions	2.1	1.9	1	2	4	6	1	44
Years Between Interactions	2.9	2.7	2	4	6	8	1	26
Patents	8.5	11.6	5	9	17	25	2	599
Distinct Firm	0.326	0.553	0	1	1	1	0	9
Distinct Region	0.049	0.235	0	0	0	0	0	5

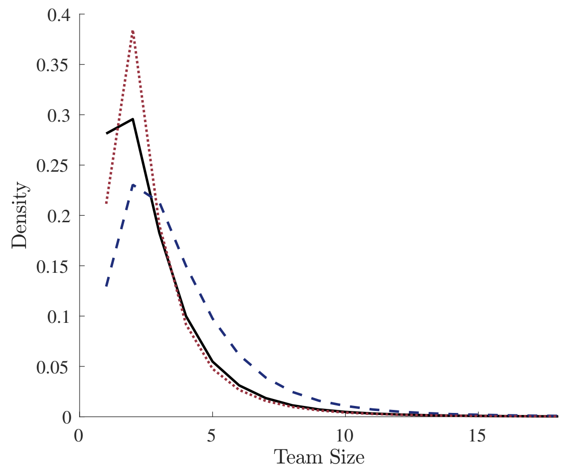
Notes: Summary stats based on 72,910 unique inventors corresponding to switcher team leaders– those who alternate between being team leaders and team members in the data after their initial role as team leaders. Panel A shows the statistics for switchers before becoming team leaders and Panel B after becoming team leaders.

A-6 Additional Figures from Section 5

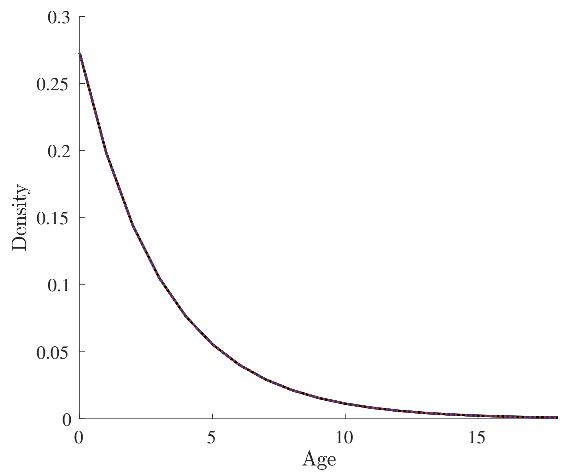
Figure A-3: Effect of Interactions and External Learning on Growth: Additional Results



(a) Productivity Growth over Inventors' Life Cycles



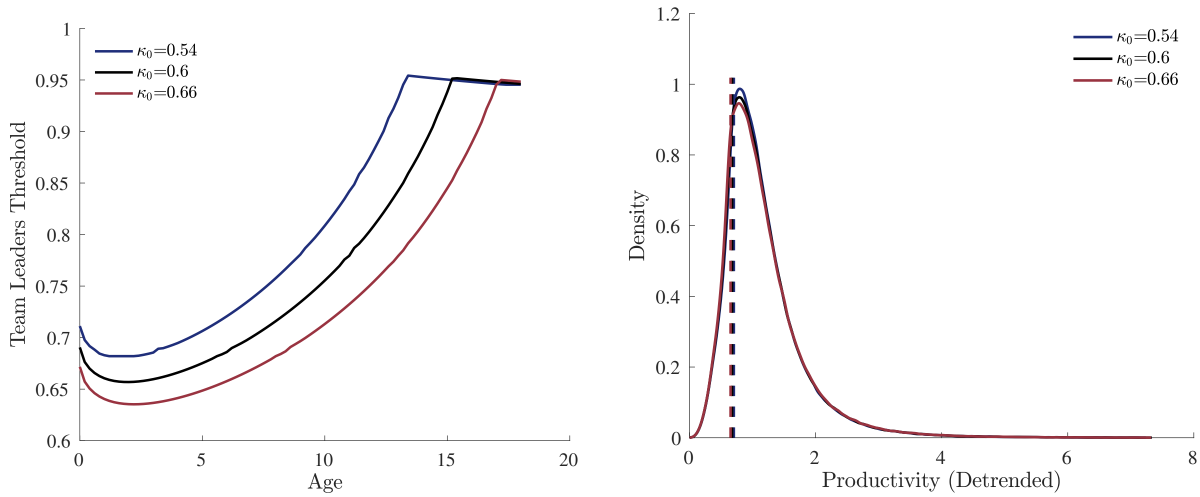
(b) Team Size Distribution



(c) Age Distribution

Panel (A) depicts average inventor productivity z across all inventors at a given age. It is normalized by productivity at age 0. Panel (B) shows the distribution of the team size n. In each panel, the solid black line represents the baseline case, with both the endogenous interaction channel and the external learning channel active. The blue dashed line corresponds to the case with no interactions within teams ($m = 0$). The red dotted line corresponds to the case with no external learning ($m_X = 0$).

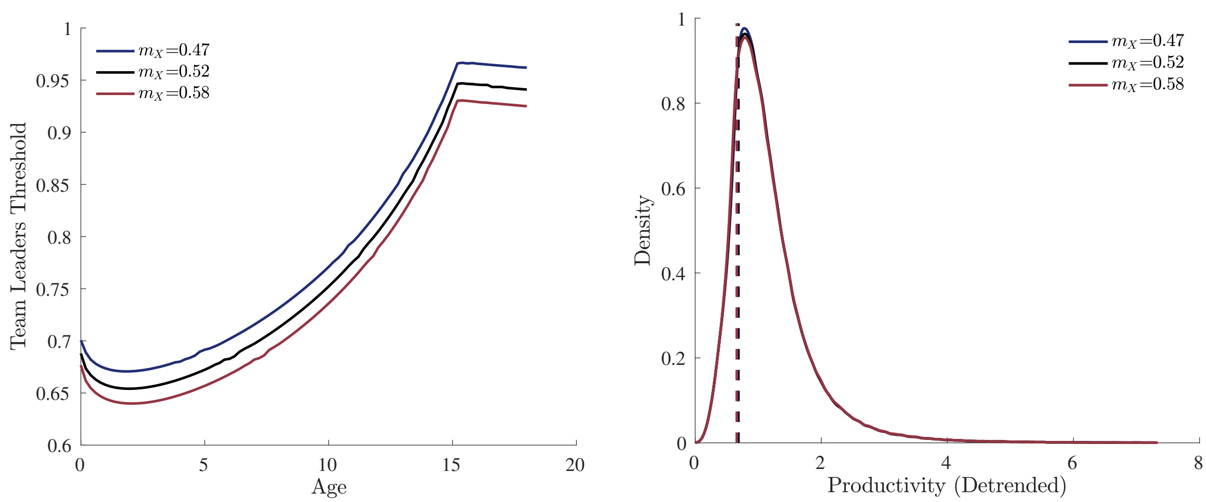
Figure A-4: Reducing Interaction Costs κ: Additional Results



(a) Team Leaders Threshold (b) Productivity Density

Panel (A) shows the threshold for being a team leader as a function of age $\hat{x}(s)$ and Panel (B): represents the distribution of de-trended idea quality $q(x) = q(z,t)e^{-g(1-\eta)t}$, for three levels of interaction costs, $\kappa_0 = 0.54, 0.6, 0.66$.

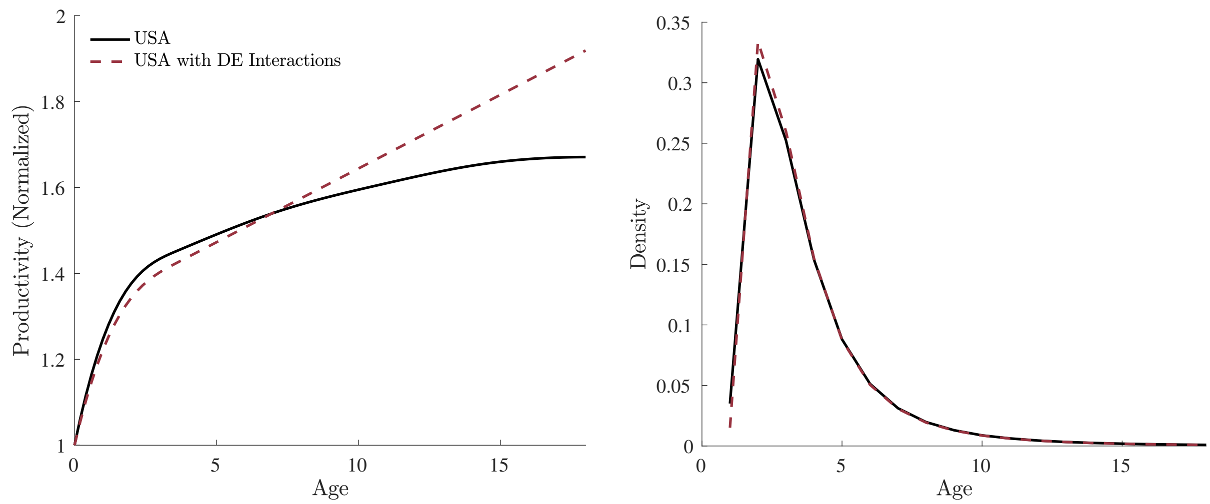
Figure A-5: Reducing the External Learning Rate: Additional Results



(a) Team Leaders Threshold (b) Productivity Density

Panel (A) shows the threshold for being a team leader as a function of age $\hat{x}(s)$ and Panel (B): represents the distribution of de-trended idea quality $q(x) = q(z,t)e^{-g(1-\eta)t}$, for three levels of external learning rate, $m_X = 0.47, 0.52, 0.58$.

Figure A-6: Reducing Interaction Costs κ_{All} in the Benchmark to Germany's Level



(a) Productivity Growth over Inventors' Life Cycles

(b) Team Size Distribution

Panel (A): represents the productivity of inventors normalized to 1 at age 0. Panel (B) depicts team size distribution. In each panel, the solid blue line represents the baseline case for the US. The black dashed line corresponds to the case changing the US interaction cost to target Germany's interactions.

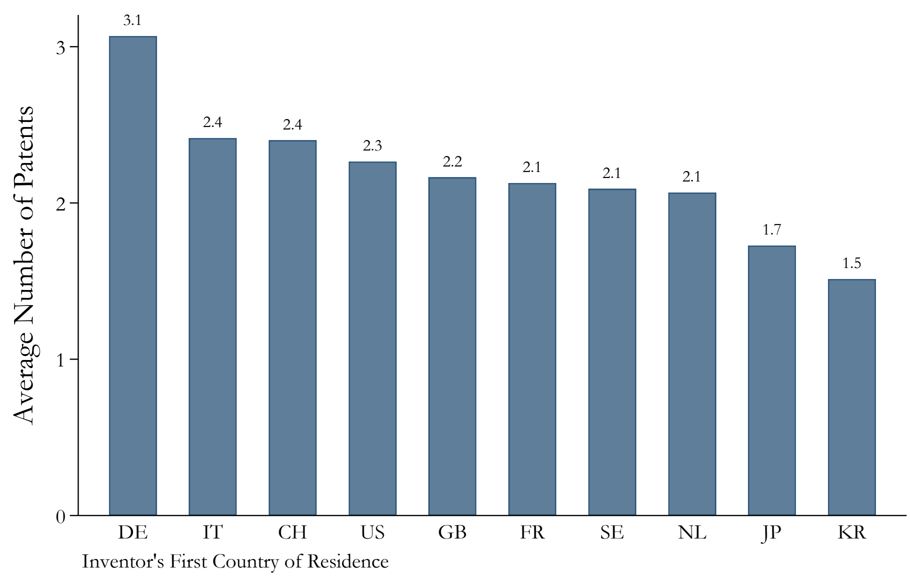
Supplement for "Dancing with the Stars: Innovation through Interactions"

Ufuk Akcigit Santiago Caicedo Ernest Miguelez
Stefanie Stantcheva Valerio Sterzi

March 27, 2025

OA-1 Data and Variable Construction

Figure OA-1: Average Number of Patents per Inventor: Top 10 Countries



OA-1.1 The Disambiguation Algorithm

For disambiguation, we rely on the Massacrator 2.0 (Pezzoni et al. (2014)). In a nutshell, Massacrator 2.0 is a disambiguation algorithm based on a 3-step procedure (see Raffo and Lhuillery (2009), for a detailed overview of inventor disambiguation approaches): first, *cleaning/parsing* of the relevant text string (inventors' name and surname, plus addresses). Second, the *matching* stage, in where the algorithm selects pairs of inventors, from different patents, who are likely candidates to be the same person, due to homonymy or quasi- homonymy of their names. Finally, the *filtering* stage, where all the pairs of inventors' names who are likely to be the same person are compared according to additional information retrieved either from the patent documentation or external sources. This includes information of common co-inventors, inventors' addresses geographical location, information on common applicants of their patents, common technological classes (according to the International Patent Classification – IPC), and the common cited prior art of the two inventors. The algorithm

Table OA-1: Number of patent applications worldwide and by selected patent offices (1995-2010)

Year	*Applications worldwide*	*Patent Families*	*Applications at EPO*	*Share EPO over applications*	*Share EPO over families*	*Applications at USPTO*	*Share USPTO over applications*	*Share USPTO over families*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1995	1,047,400	572,565	71,861	6.86%	12.55%	192,638	18.39%	33.64%
1996	1,088,400	592,413	79,104	7.27%	13.35%	194,631	17.88%	32.85%
1997	1,163,200	627,835	90,047	7.74%	14.34%	226,298	19.45%	36.04%
1998	1,214,800	665,622	102,159	8.41%	15.35%	236,149	19.44%	35.48%
1999	1,268,400	695,482	112,525	8.87%	16.18%	267,705	21.11%	38.49%
2000	1,377,400	772,496	126,947	9.22%	16.43%	321,306	23.33%	41.59%
2001	1,456,900	803,126	135,919	9.33%	16.92%	375,503	25.77%	46.76%
2002	1,443,600	806,431	134,700	9.33%	16.70%	385,722	26.72%	47.83%
2003	1,490,300	841,558	140,760	9.45%	16.73%	402,947	27.04%	47.88%
2004	1,574,400	863,716	149,644	9.50%	17.33%	453,647	28.81%	52.52%
2005	1,702,900	905,780	159,748	9.38%	17.64%	495,523	29.10%	54.71%
2006	1,791,200	929,842	166,677	9.31%	17.93%	490,291	27.37%	52.73%
2007	1,876,900	956,422	168,116	8.96%	17.58%	500,638	26.67%	52.34%
2008	1,929,200	984,022	168,860	8.75%	17.16%	488,789	25.34%	49.67%
2009	1,861,700	955,880	158,985	8.54%	16.63%	457,481	24.57%	47.86%
2010	1,996,800	1,009,086	164,431	8.23%	16.30%	478,533	23.96%	47.42%

Source: Adapted from Miguelez (2018). World Intellectual Property Indicators – 2013 edition (Wipo, 2013) and author's calculations from PatStat.

is calibrated against two benchmark datasets. The first one is a group of 530 academic French inventors, for whom all their patents were retrieved manually from PatStat and other sources. The second is a list of 342 faculty members of the EPFL ("Ecole Polytechnique Federale de Lausanne"), for whom again their patenting record is manually verified (see Lissoni et al. (2010)).

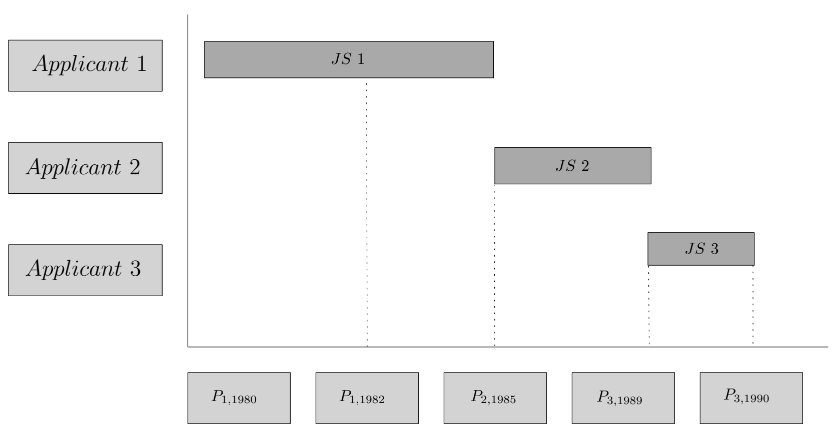
OA-1.2 Technology Fields and the International Patent Classification

The International Patent Classification (IPC) provides a hierarchical system of codes for the classification of patents and utility models according to the different areas of technology to which they pertain. This large list of technology codes can be grouped into meaningful technological fields. Here we used the classification built by the French "Observatoire des Sciences et des Techniques" (OST), which groups IPC codes into 30 meaningful fields, which are further regrouped into 7 board fields (OST, 2008). Table OA-2 shows the OST Technological classification used in the paper, as well as the distribution of patent application across fields, for the whole period under analysis.

OA-1.3 Inventors' Employment Histories

We assume that a given inventor i is active in a firm f_1 from the year of his first patent in firm f_1 to either (1) the year before the first patent was invented in firm f_2 or (2) the year of the last patent in firm f_1 in case he does not have any further patent afterward. Figure OA-2 depicts a hypothetical example of employment history construction. In that figure, the inventor is listed in five patent applications from three different applicants and has thus three job spells (JS). *JS 1* lasts from 1980, the year of the first patent in firm 1 to 1982, the year before the first patent in firm 2. *JS 2* lasts from 1983, the year for the first patent in firm 2, to 1988, the year before the first patent in firm 3. Finally, in 1989 and 1990 the inventor is in firm 3.

Figure OA-2: Example of Employment History Construction



Note: Hypothetical employment history construction for inventor listed in five patents $P_{f,year}$, from three different applicants (firms) and thus has three job spells (JS).

Table OA-2: OST Technological reclassification

30-code	30-Name	# patents	%	7-Code	7-Name
1	Electrical engineering	251,175	5.88	1	Electrical engineering; Electronics
2	Audiovisual technology	165,344	3.87	1	Electrical engineering; Electronics
3	Telecommunications	256,220	6.00	1	Electrical engineering; Electronics
4	Information technology	220,204	5.15	1	Electrical engineering; Electronics
5	Semiconductors	107,972	2.53	1	Electrical engineering; Electronics
6	Optics	136,551	3.20	2	Instruments
7	Control technology	304,393	7.12	2	Instruments
8	Medical engineering	195,726	4.58	2	Instruments
9	Nuclear technology	18,774	0.44	2	Instruments
10	Organic chemistry	195,457	4.58	3	Chemicals; Materials
11	Macromolecular chemistry	161,291	3.78	3	Chemicals; Materials
12	Basic chemistry	146,652	3.43	3	Chemicals; Materials
13	Surface technology	95,226	2.23	3	Chemicals; Materials
14	Materials; Metallurgy	101,127	2.37	3	Chemicals; Materials
15	Biotechnologies	151,143	3.54	4	Pharmaceuticals; Biotechnology
16	Pharmaceuticals; Cosmetics	252,034	5.90	4	Pharmaceuticals; Biotechnology
17	Agricultural and food products	45,034	1.05	4	Pharmaceuticals; Biotechnology
18	Mechanical engineering (excl. Transport)	159,877	3.74	5	Industrial processes
19	Handling; Printing	178,495	4.18	5	Industrial processes
20	Agricultural and food apparatuses	168,358	3.94	5	Industrial processes
21	Materials processing	38,158	0.89	5	Industrial processes
22	Environmental technologies	42,269	0.99	5	Industrial processes
23	Machine tools	94,044	2.20	6	Mechanical eng.; Machines; Transport
24	Engines; Pumps; Turbines	109,656	2.57	6	Mechanical eng.; Machines; Transport
25	Thermal processes	63,544	1.49	6	Mechanical eng.; Machines; Transport
26	Mechanical elements	144,345	3.38	6	Mechanical eng.; Machines; Transport
27	Transport technology	175,345	4.10	6	Mechanical eng.; Machines; Transport
28	Space technology; Weapons	15,523	0.36	6	Mechanical eng.; Machines; Transport
29	Consumer goods	166,739	3.90	7	Consumer goods; Civil engineering
30	Civil engineering	111,541	2.61	7	Consumer goods; Civil engineering

Notes: Patents assigned to more than one category are counted more than once.

We make use of the HAN database (OECD HAN database, January 2014) of Harmonized Applicants' Names in order to track assignees over time.

The OECD HAN database, January 2014 edition, combines information on 317,927 applicants' names listed in 2,776,774 patent applications. Of them, 2,519,069 (90.7%) have only one applicant, while 9.3% patents list more than one. For the remaining 178,281 patent applications with no link to the HAN database, harmonized applicants' names are extracted from the CRIOS-PatStat database (Coffano and Tarasconi (2014)).

OA-2 Computational Algorithm

In this appendix we describe the details of the simulation of the model. We simulate the model for $n_{sim} = 100,000$ individuals for $T = 25$ years, and subdivide each period using a step size of $\Delta t = 0.2$. Additionally, we describe the algorithm we use to solve the system of nonlinear equations simultaneously. This allows us to have a distribution of productivity $\Phi(x)$ consistent with an $\hat{x}(s)$ function that solves the HJB equations (A-5) for all s and with w that clears the market as in (A-6).

Model Simulation

In this section, we describe the process to simulate the model. Given $\hat{x}(s)$ and a set of parameters, this process yields the distribution of productivity and a record of all individuals' interactions across their lifespan.

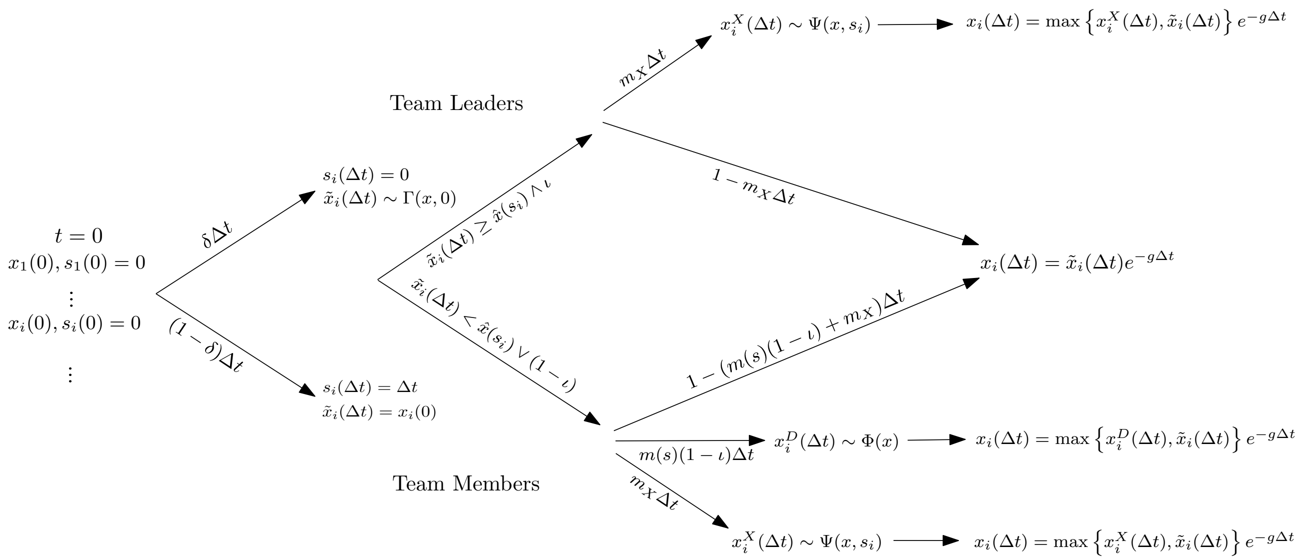
1. Every agent begins with an initial draw from a Log-logistic distribution $\Gamma(x, 0) = \frac{1}{1+k_0 x^{-1/\theta}}$. To simulate the initial draw, we use the inverse distribution method.

2. With probability δ agents die. If they do die, it is instantly replaced by an age 0 agent with a productivity draw $x \sim \Gamma(x, 0)$. Otherwise, their age change by Δt, and one of the following happens.

   - Agents have differential learning processes depending on their age and their productivity. Let $\tilde{x}(t + \Delta t) = x(t)$ be the productivity of an agent before learning.
   - If $\tilde{x}(t + \Delta t) \geq \hat{x}(s(t))$ then one of the following happens:
     (a) With probability $m_X \Delta t$ the agent gets a draw from the external source $x^X(t + \Delta t) \sim \Psi(x, s(t + \Delta t))$. Then the productivity of the agent is $x(t + \Delta t) = \max\{x^X(t + \Delta t), \tilde{x}(t + \Delta t)\}e^{-g\Delta t}$.
     (b) With probability $(1 - \iota)m(s)\Delta t$ the agent draws a from the endogenous productivity distribution $x^D(t + \Delta t) \sim \Phi(x)$. Then the productivity of the agent is $x(t + \Delta t) = \max\{x^D(t + \Delta t), \tilde{x}(t + \Delta t)\}e^{-g\Delta t}$.
     (c) With probability $1 - (m(s)(1 - \iota) + m_X)\Delta t$ the agent does not learn and thus $x(t + \Delta t) = \tilde{x}(t + \Delta t)e^{-g\Delta t}$.
   - If $\tilde{x}(t + \Delta t) < \hat{x}(s(t))$ then one of the following happens:
     (a) With probability $m_X \Delta t$ the agent gets a draw from the external source $x^X(t + \Delta t) \sim \Psi(x, s(t + \Delta t))$. Then the productivity of the agent is $x(t + \Delta t) = \max\{x^X(t + \Delta t), \tilde{x}(t + \Delta t)\}e^{-g\Delta t}$.
     (b) With probability $m(s)\Delta t$, the agent gets a draw from the productivity distribution from team leaders $x^D(t + \Delta t) \sim \Phi(x)$. Then the productivity of the agent is $x(t + \Delta t) = \max\{x^D(t + \Delta t), \tilde{x}(t + \Delta t)\}e^{-g\Delta t}$.
     (c) With probability $1 - (m(s) + m_X)\Delta t$ the agent does not learn and thus $x(t + \Delta t) = \tilde{x}(t + \Delta t)e^{-g\Delta t}$.

The full diagram of the simulation is presented in Figure OA-3.

FIGURE OA-3: SIMULATION PROCEDURE



Solving the System of Equations

In this section, we describe in detail the algorithm used to solve the system of equations. The difficulty of solving the model is due to the fact that $\hat{x}(s)$ is an input to simulate the economy as described before and obtain the endogenous distribution of productivity $\Phi(x)$ but also is the result of solving a system of equations that takes $\Phi(x)$ as a given input.

1. We begin with an educated guess of $\Phi_0(x)$, g_0 and solve for $\hat{x}(s)$ and w. [55]

   (a) Begin with a guess of w_0. Using a similar procedure to (Lucas and Moll, 2014), we solve (3) using finite differences to obtain $v_\ell(x, s)$ for each s. We can solve the system of equations numerically by discretizing the problem and using the first-difference method.

   - Discretize the productivity space into I points, $\{x_1, ..., x_I\}$
   - Let $h_i = x_i - x_{i-1}$ and approximate the derivative using the backward difference, $v'(x_i) \approx \frac{v_i - v_{i-1}}{h_i}$

   Team Members

   - Team members of age s,

$$(\rho + \delta - g(1-\eta))v_i^m + g\frac{v_i^m - v_{i-1}^m}{h_i}x_i = w_0 + m\sum_{j=i}^{I}\left[\iota v_j^l + (1-\iota)v_j^m - v_i^m\right]\phi_j h_j$$
$$+ m_X\sum_{j=i}^{I}\left[\iota v_j^l + (1-\iota)v_j^m - v_i^m\right]\psi_j h_j$$
$$+ \iota(v_i^l - v_i^m) \;\; \forall i \geq 1$$

[55]To obtain an educated guess we solve the model assuming that $\hat{x}(s) = \bar{x}$ for every s.

- Rearranging the terms, we can write the system of equations as,

$$a_i^m v_i^m + b_i^m v_{i-1}^m + \sum_{j=i+1}^{I} c_j^m v_j^m + a_i^l v_i^l + \sum_{j=i+1}^{I} c_j^l v_j^l = w_0 \tag{OA-1}$$

where,
  - $a_i^m = \rho + \delta - g(1-\eta) + \iota - m(1-\iota)\phi_i h_i + m(1-\Phi_i) - m_X(1-\iota)\psi_i h_i + m_X(1-\Psi_i) + g\frac{x_i}{h_i}$
  - $b_i^m = -\frac{g x_i}{h_i}$
  - $c_j^m = -m(1-\iota)\phi_j h_j - m_X(1-\iota)\psi_j h_j$
  - $a_i^l = -\iota - m\iota\phi_i h_i - m_X \iota\psi_i h_i$
  - $c_j^l = -m\iota\phi_j h_j - m_X\iota\psi_j h_j$

Team Leaders

- Team leaders of age s,

$$(\rho + \delta - g(1-\eta))v_i^l + g\frac{v_i^l - v_{i-1}^l}{h_i} x_i = \pi_0 x_i + m_X \sum_{j=i}^{I} \left[\iota v_j^l + (1-\iota)v_j^m - v_i^m\right] \psi_j h_j + (1-\iota)(v_i^m - v_i^l) \quad \forall i \geq I+1$$

- Rearranging the terms, we can write the system of equations as,

$$a_i^m v_i^m + b_i^m v_{i-1}^m + \sum_{j=i+1}^{I} c_j^m v_j^m + a_i^l v_i^l + \sum_{j=i+1}^{I} c_j^l v_j^l = \pi_0 x_i \tag{OA-2}$$

where,
  - $a_i^m = -(1-\iota) - m_X(1-\iota)\psi_i h_i$
  - $c_j^m = -m_X(1-\iota)\psi_j h_j$
  - $a_i^l = \rho + \delta - g(1-\eta) + (1-\iota) - m_X(1-\iota)\psi_i h_i + m_X(1-\Psi_i) + g\frac{x_i}{h_i}$
  - $b_i^l = -\frac{g x_i}{h_i}$
  - $c_j^l = -m_X \iota \psi_j h_j$
- Note that stacking equations (OA-1) and (OA-1) this is a system of $2I$ linear equations on (v_i^m, v_i^l)
- We write the system in matrix form as $Av = b$,

$$A = \begin{bmatrix} A_1^m & A_1^l \\ A_2^m & A_2^l \end{bmatrix} \; ; \; v = \begin{pmatrix} v_1^m \\ \vdots \\ v_I^m \\ v_1^l \\ \vdots \\ v_I^l \end{pmatrix} \; ; \; b = \begin{pmatrix} w_0 \\ \vdots \\ w_0 \\ \pi_0 x_1 \\ \vdots \\ \pi_0 x_I \end{pmatrix}$$

where,

$$A_1^m = \begin{pmatrix} a_1^m & c_2^m & c_3^m & \cdots & c_I^m \\ b_2^m & a_2^m & c_3^m & \cdots & c_I^m \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & c_I^m \\ 0 & \cdots & 0 & b_I^m & a_I^m \end{pmatrix} ; \; A_1^l = \begin{pmatrix} a_1^l & c_2^l & c_3^l & \cdots & c_I^l \\ 0 & a_2^l & c_3^l & \cdots & c_I^l \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & c_I^l \\ 0 & \cdots & 0 & 0 & a_I^l \end{pmatrix}$$

$$A_2^m = \begin{pmatrix} a_{I+1}^m & c_{I+2}^m & \cdots & c_{2I}^m \\ 0 & a_{I+2}^m & \cdots & c_{2I}^m \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & & \vdots \\ \vdots & \ddots & & c_{2I}^m \\ 0 & \cdots & 0 & a_{2I}^m \end{pmatrix} ; \; A_2^l = \begin{pmatrix} a_{I+1}^l & c_{I+2}^l & \cdots & c_{2I}^l \\ b_{I+2}^l & a_{I+2}^l & \cdots & c_{2I}^l \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & & \vdots \\ \vdots & \ddots & & c_{2I}^l \\ 0 & \cdots & b_{2I}^l & a_{2I}^l \end{pmatrix}$$

(b) Using the indifference condition (A-5), solve $\hat{x}(s)$ for each s from (A-4).

(c) Having estimated $\hat{x}(s)$, update the guess of w_i by solving (A-6).

(d) Keep updating $\hat{x}(s)$ and w until $\|w_{i-1} - w_i\| < \varepsilon$.

2. Using the estimated $\hat{x}(s)$, we simulate the economy using the algorithm described in OA-3. This yields a new distribution of productivity $\Phi(x)$ and a new growth rate g.

3. Repeat step 1 and 2 until $\|\Phi_i(x) - \Phi_{i-1}(x)\| < \varepsilon$ where

$$\|\Phi_1(x) - \Phi_2(x)\| = \frac{1}{100} \sum_{p \in P} |\Phi_1^{-1}(p) - \Phi_2^{-1}(p)|, \tag{OA-3}$$

with $P = \{0.01, 0.02, ..., 0.99, 1\}$.

The algorithm is computationally demanding. To speed up the computations, we parallelize the $\hat{x}(s)$ estimation for each realization of s. Additionally, we only estimate $\hat{x}(s)$ between 0 and age 18 since the mass of agents outside this interval is minimal.