Idea Production and Team Structure

Jeremy Pearce*

University of Chicago

May 17, 2022

Abstract

In the modern world, teamwork has become the bedrock of innovation. This paper builds a model that connects idea production in teams to aggregate innovation to quantify the sources and policy implications of the rise of the importance of teams in innovation. I find that most of this rise is due to the changing production function of ideas, implying that policies that interact with innovation must address the team production channel. With the calibrated model, I ask how aggregate innovation responds to exogenous shifts in taxes and immigration. I find that taxes have significant effects on innovation even with fixed labor supply through changing the sorting pattern in teams, pointing to a new mechanism governing the interaction of taxation and innovation. Using the fall of the Soviet Union, I show how opportunities in teams induce self-selection in migration, helping explain the outsized role of immigrants in American innovation.

Keywords: Innovation, Human Capital, Teams, Taxation, Immigration, Endogenous Growth.

JEL Classification: O30, O31, O32, J22, J61.

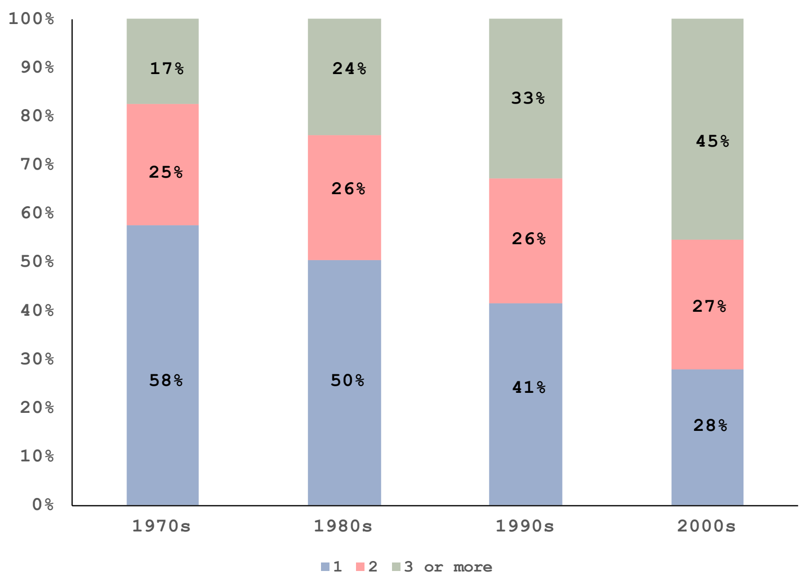
---

*I am deeply indebted to Ufuk Akcigit, Steve Davis, and Nancy Stokey, for tremendous insight and support. Eliot Abrams, Stephane Bonhomme, Thomas Bournay, Santiago Caicedo, Michele Carter, Mark Goldsmith, John Grigsby, Ben Jones, Greg Kaplan, Bill Kerr, Ilse Lindenlaub, Josh Morris-Levinson, Karthik Nagarajan, Marta Prato, Esteban Rossi-Hansberg, Gustavo de Souza, Stefanie Stantcheva, Stratton Tolmie, Alessandra Voena, Liangjie Wu, and Yulia Zhestkova have all provided valuable and insightful comments, as have participants in seminars at the University of Chicago, Harvard Business School, and CREI. Contact: `jgpearce@uchicago.edu`.

1 Introduction

Today's technological and scientific innovations rely heavily on teamwork, even in traditionally individualistic fields. According to mathematician Terence Tao, "Nowadays, most problems in mathematics are interdisciplinary. You need expertise from many different fields of mathematics or subfields outside mathematics. More and more, you need to collaborate".[1] Figure 1 illustrates this increasing reliance on teamwork over time; among patents registered with the U.S. Patent and Trademark Office (USPTO), the fraction of sole-inventor patents has fallen from around 60% of patents in the 1970s to less than 30% of patents in the first decade this century.[2] Meanwhile, teams have also become more diverse across ethnicity, gender, and field background.[3] Teams of inventors with heterogeneous backgrounds and expertise have become central to innovation.

Figure 1: Team size by decade on USPTO patents



Notes: This figure plots the shares of number of authors listed on a patent (e.g. team size on an invention). The shares plotted are one, two, or three or more respectively by decade in the US. Source: USPTO.

Innovation is the central determinant of long-run economic growth. Thus, as teams become a more central component of modern innovation, they become more essential for sustaining long-run growth. Further, different explanations for this rise in teams have different implications for

[1] Tao (2017).
[2] This fact has been documented in Wuchty et al. (2007) in both patents and scientific publications.
[3] See Appendix B.2 for these three facts.

many policies such as taxation, immigration, education. However, there is not a comprehensive framework to analyze the links between teams of differentiated skills and aggregate innovation.

This paper investigates and quantifies the sources and implications of the increasing importance of teams in innovation. In quantifying the sources of the rise of teams, I focus on three classic economic concepts: *benefits*, *costs*, and the *inventor supply*. I map these three concepts into data-relevant counterparts: changes in the idea production function, which takes as inputs the *depth* and *breadth* of team expertise and outputs an idea of a given quality (*benefits*); reductions in communication costs (*costs*); and changes in the aggregate composition of inventor expertise (*inventor supply*). The quantification procedure admits a range of policy counterfactuals. I focus on taxation and immigration policy and use the fall of the Soviet Union to verify the model framework and study the immigration application.

The paper proceeds in four steps. First, I build a matching-in-teams model that links the benefits and costs of a specific team to how often that team forms. The model allows for a quantitative decomposition of the roles of the three forces (benefits, costs, supply) in determining the changing pattern of teams. The model follows most closely classic marriage models but allows for matching across teams of multiple sizes of any type (i.e., not restricted to two-sided markets).

Second, I use USPTO patent data to build measures of individual skill expertise that are based on the distribution of an individual's work across patent classes that allows comparison of skills and teams over time. The measures admit a parsimonious connection between the model ingredients of expertise types with differentiated skills and the data ingredients of differentiated experience of individuals across technologies.

Third, I connect the model to the skill measurement to uncover patterns in team production. I quantify the contribution of the three forces by embedding the estimated production function and communication costs into the matching model. I compare the 1980s to the 1990s when evaluating the rise in team size. While this change is part of a larger trend, by restricting attention to the last decades of the 20th century, I deal with no issues on the measurement of the quality of patents. I find that changes in the benefits to teams or the idea production function explain the largest share of the increase in team size (68%). Falling communication costs across geographical distance also explains a significant portion of the rise in teams (25%). Another potential source of the rise of teams is the composition of inventors, e.g. supply. If chemists work in large teams and chemistry becomes a more innovative field, then team size would increase mechanically. Yet in the quantitative exercise, I find a negative effect of composition on the change in team size, contributing -9% to the observed increase.

Fourth, I explore policy applications within the team production and matching framework, with a particular focus on taxation and immigration policy. The team economy provides a new mechanism to think through how policies and aggregate innovation interact. Taxes and immigration address two different margins on the interaction of policies and innovation. First, taxes conditional on an existing distribution of skill focuses purely on how taxation affects sorting into innovative teams. Second, exogenous labor supply shifts show how H1B policies can affect innovation. The model provides a way to think through how skill-targeted policies interact with the economy. These two different margins address how inventors sort into teams and the overall skill composition in the economy.

In an application of this policy exercise, I make use of quasi-experimental evidence from the breakup of the Soviet Union. Russians[4] consistently comprised around 0.6% of inventors in the US in until the early 1990s, but accounted for 1.4% of inventors by 2005. Leveraging the rise of Russian inventors, I shock the set of skills in the United States with the Russian skills. To do this, I use patent records from the Soviet Union that illustrate how Russian skills across technologies differed from US skills across technologies in the years before the breakup in 1991. This inventor expertise shock exercise finds that the self-selection in migration and the patterns of Russians across teams can be predicted by the model. I also use this evidence to evaluate the contribution of Russians after the fall of the Soviet Union and find that the production in teams explains a significant component of their impact on US innovation. Without a model of production and matching in teams, it would not be possible to make a statement on the quantity of their overall contribution.

Related Literature

Economists have long known that ideas are the building blocks of long-run economic growth (Lucas, 1988; Romer, 1990; Aghion and Howitt, 1992) and modern endogenous growth frameworks have moved to put human capital at the center (Akcigit et al., 2020). Economists have also noted and hypothesized about the rise of teams in idea production (Wuchty et al., 2007). While there are

[4] I use ethnicities associated with the former Soviet Union as in Kerr (2007).

competing theories on this rise (Jones, 2009; Bloom et al., 2020), the connection between specific expertise in teams and aggregate innovation has had little exploration. Given the rise of teams, the importance of heterogeneous expertise in innovation is a natural issue to command the attention of researchers in innovation and economic growth.

Two common hypotheses exist in explaining the rise of teams: changes in the expertise required to produce new ideas (the benefits to teams – Falk-Krzesinski et al., 2010; Bennett and Gadlin, 2012) and changes in communication costs (Adams et al., 2005; Agrawal and Goldfarb, 2008; Jones et al., 2008; Forman and van Zeebroeck, 2012). Although there is a rising interest among growth economists in team mechanics (Akcigit et al., 2018) and the interaction with technology (Porzio, 2017), the role of horizontally differentiated expertise has not been incorporated into the growth literature. Given that human capital is the key component driving innovation and long-run growth (Barro, 2001; Ciccone and Papaioannou, 2009; Waldinger, 2016), understanding this horizontal expertise will provide an important lens for thinking about how the distribution of innovative skills affects innovation.

To understand the rise of teams and the implications, I build on two literatures: i) endogenous economic growth and ii) matching and production in teams. In the case of empirical endogenous growth, there is a significant literature documenting the connection between patents and innovation. Akcigit et al. (2017b) find that regions and economies that produce more patents grow faster over the long-run. Bloom and Van Reenen (2002), Kogan et al. (2017), and Akcigit and Kerr (2018) show patents have a significant impact on firm-level productivity, stock market value, and employment. In addition to creative destruction and economic growth, patents also transfer knowledge. Jaffe and Trajtenberg (1999) illustrate the role that patents play in facilitating flows of international knowledge and Bloom et al. (2013) focus on the interaction of technology spillovers and the product market. Henderson et al. (1998) and Ahmadpoor and Jones (2017) further show that patents channel basic academic research to applied research and then to products.

Theoretical economists have stressed the connection between the production of ideas and aggregate innovation (e.g., Jones, 2005; Lucas, 2009). Lucas and Moll (2014) and Perla and Tonetti (2014) focus on the choice to invest in learning versus production in driving aggregate innovation, while Benhabib et al. (2014) and Konig et al. (2016) focus on the role of imitation versus innovation at firms, which also happens at the individual level. Eeckhout and Jovanovic (2002) address how imitation and knowledge spillovers can interact with the determinants of inequality. Luttmer (2015) focuses on this inequality and the endogenous matching process that creates students and teachers. Further, there have been empirical and theoretical frameworks providing evidence that skill breadth and depth and locational distribution matter for idea quality (e.g. Jones, 2009; Berkes and Gaetani, 2019). While idea heterogeneity serves an important role in the endogenous growth literature, its roots in domain specific human capital and individual interaction have received less attention. Yet, heterogeneous human capital is at the core of heterogeneous ideas.

This paper puts heterogeneous human capital at the center. I embed this heterogeneity into a matching model wherein individuals with differentiated expertise productivity depends on the team they join. The theoretical framework links the distribution of expertise to matching patterns in teams, building on the framework Becker (1973) developed to study marriage. The paper additionally includes insights on the division of labor with skill heterogeneity (Becker and Murphy, 1992; Stokey, 2018). I stress the importance of skill both being multidimensional (Lindenlaub, 2017) and specific to certain knowledge domains (Hayek, 1945). The organization of an economy and the teams within it depends on this distribution of knowledge (Garicano and Rossi-Hansberg, 2006). In building the interaction of heterogeneous knowledge in teams into the matching process, the model most closely follows Choo and Siow (2006), who estimate a marriage model with transferable utility and heterogeneous preferences following Becker (1973) and McFadden (1974). There has been recent work on extending this empirical framework to imperfectly transferable utility (Galichon et al., 2019). This current paper focuses on teams where the same type can match with each other and transferable utility ensures market clearing.

This paper also builds on an empirical literature on the mechanics of team production. I direct significant focus in this paper to diverse expertise in teams which economists have discussed as increasing patent impact (Singh and Fleming, 2010). Certain studies find positive (Jehn, 1995, 1997) and negative effects (De Dreu and Weingart, 2003) of diversity (or horizontal differentiation) on team production. Many papers find a non-monotonic effect (e.g. Guimera et al., 2005; De Dreu, 2006; de Wit et al., 2011). In this paper, I argue that there is a key connection between the underlying expertise and the technology that will determine the payoffs to different expertises working together. Furthermore, results from surprise deaths indicate that team complementarities are significant (Azoulay et al., 2010; Jaravel et al., 2018). Previous work has also suggested this element of team complementarities is carrying increasing importance over time (Freeman et al., 2014). This current paper explores the matching process underlying innovation to understand these forces.

The rise of sorting into teams has motivated academics to further investigate the foundational reasons for forming teams (Bikard et al., 2015; Teodoridis, 2018; Wu et al., 2019). Even though the split of the returns to innovation may have significant heterogeneity across inventors within a team (Kline et al., 2019), the rise of teamwork suggests the benefits are increasing. Indeed, there is microeconomic theory on the benefits of cognitively diverse teams (Hong and Page, 2004; Page, 2007) and how the response of teams to skill determines wages (Davis, 1997). The goal of this paper is to bring the intuition of these studies in a general equilibrium environment with innovation.

Disciplining the role of expertise in teams and innovation has many potential applications. This paper primarily directs its attention to taxation and immigration policies. I build on a literature that delivers conflicting accounts of the role played by immigration in idea production. In one study, Borjas and Doran (2012) find that Russian mathematician immigrants substituted for US mathematicians in idea production in mathematics. In another study, Moser et al. (2014) find positive spillovers from German-Jewish chemist emigres. Historically, immigrants have had a significant impact on American technology (Akcigit et al., 2017a). Immigrants are more concentrated in patent-heavy fields, leading them to be more innovative than natives (Hunt and Gauthier-Loiselle, 2010). Given the significant moving costs, it makes sense that immigrants who move self-select into productive fields (Borjas, 1987). In regards to policy, Kerr and Lincoln (2010) find that increases in the H1B cap spurred innovation. In this paper, I stress the team production channel is essential for understanding how immigrants interact with the countries' existing expertise to shape overall output. I find that the geographical distribution of individuals still has first-order implications on the prevailing team structure and economic growth (Porter and Stern, 2001), which suggests policies related to immigration continue to be an essential ingredient in innovation. Addressing horizontal skill differentiation is crucial in exploring the interaction between immigrant and native expertise.

The paper is organized as follows. Section 2 describes three empirical facts that motivate a quantitative model. Section 3 builds a team production model that links the value of teams to their observed frequency. Section 4 introduces the USPTO data and the construction of measures I use in the empirical analysis. Section 5 discusses the quantitative decomposition of the role of each force in driving the change in team size. Section 6 illustrates the ability of these results to elucidate policy questions and focuses on immigration policy. Section 7 discusses the general robustness of the results from the previous sections. Section 8 concludes.

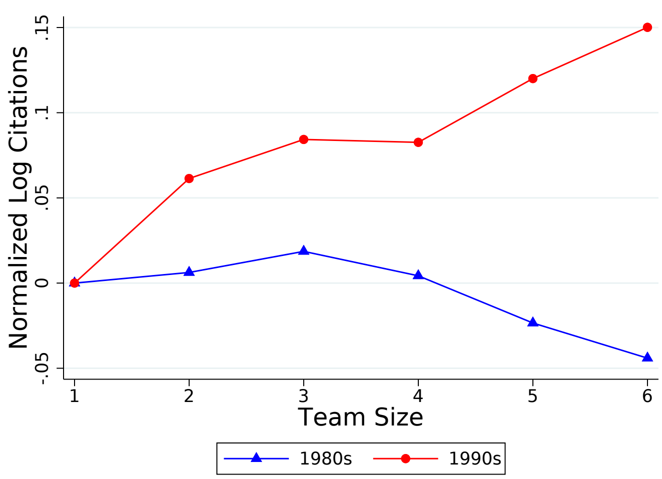
2 Three Facts on Team Production in R&D

This section presents three facts related to changing forces in idea production in teams: *benefits*, *costs*, and *supply*. These three forces will serve as an important reference point for the analysis for this paper. I use USPTO patent data that has the listed inventors and their residences.

Fact 1: The impact of team patents has been rising over time.

Teams increasingly produce higher cited patents than individuals. Figure 2 plots log lifetime citations on a patent by team size during the 1980s versus the 1990s.[5] During the 1990s, teams larger than two, relative to individuals working alone, created higher impact patents than did teams of similar size during the 1980s.

Figure 2: Log Lifetime Citations by Team Size



Notes: This figure plots the average log citations normalized by the single-author value in the 1980s and the 1990s. Thus, each dot represents the average log citations within each team size relative to sole-authored patents. Source: USPTO.

[5] Log lifetime citations takes the current number of citations on a patent and adjusts by patent age and class. This adjustment accounts for the fact that older patents tend to have more citations. These citations proxy for the overall social value of a patent.

The 1990s exhibit stronger growth in expected log citations to team size than the 1980s, which shows little response. This suggests that the *benefits* to teams have risen. This is also true for output measured by stock market value as in Kogan et al. (2017).[6] Over time, combining individuals together has led to relatively more productive innovations.[7] This is true also when controlling for the amount of time spent on a patent.[8] The relatively higher output of teams suggests that in order to make use of this increased output, inventors would respond by working relatively more often in teams than they work alone. This highlights the changing benefit of teams.

Fact 2: Team collaborations are becoming more geographically dispersed over time. In addition to the output effect, a cost effect could explain the rise of teams. Individuals may form teams more often because team members can more easily communicate across geographical distances (e.g., email, internet, phone, travel costs). This has been studied from a theoretical angle as a key feature of the division of labor (e.g. Grossman and Rossi-Hansberg, 2008; Costinot, 2009). Figure 3 plots the proportion of two-person teams that worked together while at least 100 miles apart from 1980-2005. If teams with more than two inventors are included, the pattern becomes stronger. The increased frequency of geographically distant teams is suggestive of falling geographical communication costs, and a lower cost of forming teams.

Fact 3: Immigrant inventors have made significant contributions to computer technology. Finally, I examine how the composition of inventors across technologies can shift the distribution of teams. Many immigrants arrived in the US in the 1990s and produced patents in technical fields. Figure 4 focuses specifically on Russians. Using ethnicity name probability matches, Figure 4a shows that ethnic Russians represented consistently around 0.6-0.7% of inventors in the 1980s. In the 1990s, this percentage increased sharply, rising to around 1.4% of US inventors by 2005.

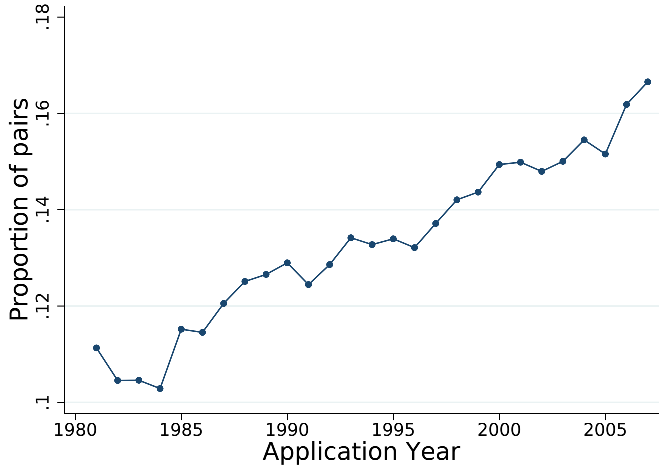
Figure 4b illustrates that inventors from the former Soviet Union brought with them expertise that differed from the US population. Russians were especially strong in physics (with computing-related fields a subset) and construction. This figure uses patent records from the former Soviet Union in order to characterize the expertise of the population before the fall of the Soviet Union

[6] See Appendix B.1 for robustness of this result.

[7] This has been noted in scientific publications by Wuchty et al. (2007).

[8] This fact can be seen in Appendix Figure B.7.

Figure 3: Proportion of co-authors in a different location



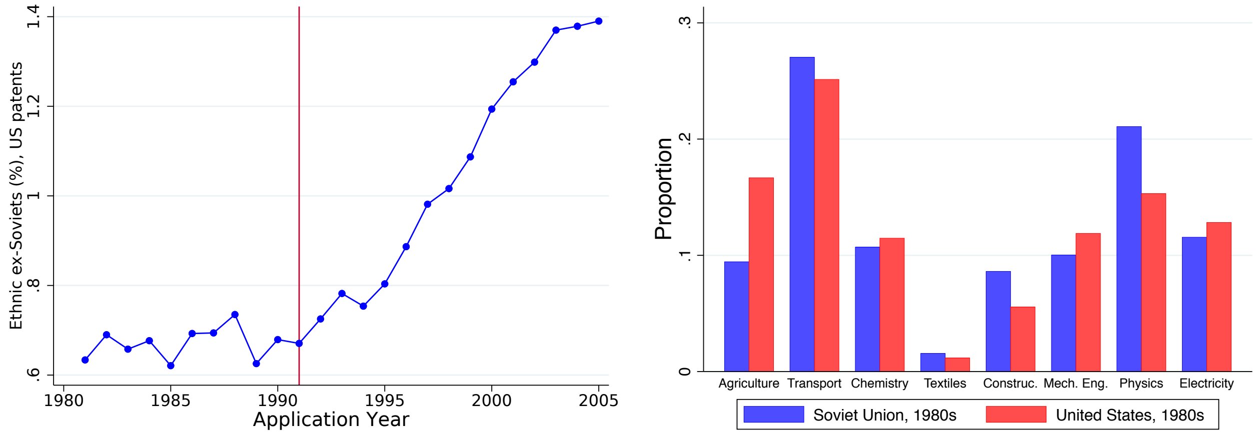
Notes: This figure plots the share of two-person co-authors in a different location over time from 1981-2006. Source: USPTO.

in 1991. When Russians came to the US they were more concentrated in patent classes that were heavily represented in the Soviet Union. This can be seen in the ethnic composition of inventors on patents in the US.

This immigration inflow was part of a larger phenomenon of immigrant influx from Russia, China, and India that shifted the expertise distribution of US inventors over time. If these inventors were more concentrated in team-intensive technologies, the composition of inventors alone could shift the average observed team size in the economy.

The main facts illustrate three key elements for the analysis of this paper. First, the benefits to teams are changing as teams produce relatively more impactful patents over time. Second, the costs of forming teams at a distance has fallen, which has induced more inter-regional collaboration. Third, immigrants are self-selected and shaping the skill composition of inventors. These elements are explored further in the following sections. I now present a model that will serve as a foundation for the quantitative analysis.

Figure 4: Immigrants and Technologies



(a) Russians on US Patents

(b) Soviet and US Tech Dist (IPC1)

Notes: Panel (a) plots the share of ex-Soviet bloc ethnicities on US patents in the USPTO. Panel (b) shows the technological distribution of patents in the Soviet Union compared to patents in the US in the 1980s. Source: USPTO, FIPS and Kerr (2007).

3 A Model of Idea Production and Team Formation

To understand the market for invention, I build a model of inventors who match in teams, produce ideas, and share the returns. Inventors observe the set of possible teams they can join and make a decision to join based on the payoffs. This model embeds the intuition of the three forces discussed in the previous section that determine the matching pattern: the benefits to teams, the costs of team formation and communication, and the supply of expertise.

The model follows a similar structure to Choo and Siow (2006), who use matching pairs in a two-sided framework to infer the value of the marriage. Here, the model is extended to a situation in which agents can take any role in a team; that is, agents can match with any other inventor and in multi-inventor teams. This fits the market for invention, as inventors often match in large teams and match with inventors of the same "type" as themselves (e.g. chemists work with chemists).

3.1 Environment

There are a mass of inventors $M_{x,\ell}$, where each individual is one of a discrete number of skill types and one of a discrete number of locations. A skill type is indexed by $x \in \mathbb{X} \subset \mathbb{R}^S$ and location indexed by $\ell \in \mathbb{L}$. While x is one of a finite number of types, x contains a vector of length S

which signifies the S domains of expertise that can be utilized to produce patents. ℓ is the single geographical location of each agent with the set $\mathbb{L}$ of total locations. The skills of team members matter for the quality of patent production. Geography matters for the costs of communicating in teams.

The inventors are risk-neutral and maximize linear utility in their wage and an idiosyncratic preference shock in a static setting. Inventors can either join a team up to some size $\overline{T}$ or work alone. Due to the finite team size $\overline{T}$ and finite number of types (x, ℓ), there are a finite number of total team types.

A team is indexed by $k = \{T; (x^1, \ell^1), ...(x^T, \ell^T)\}$. Team k generates output that is a function of the vector of skills of each type on the patent. The production functions for each team and individuals working alone are as follows:[9]

$$q_k = q(x^1, ..., x^T) \qquad ; \qquad q_{\underline{k}}(x) = q(x^1) \tag{1}$$

Each operating team needs to pay a cost to communicate that depends on the team size and geographical dispersion of inventors, c_k. This cost can be understood as the cost of communicating on a project or the cost of forming the team. The communication cost for a single inventor working alone is zero; for multi-person teams, communication costs would likely increase with the geographical dispersion of the types and the number of team members:

$$c_k = c(\ell^1, .., \ell^T) \qquad ; \qquad c_{\underline{k}} = 0 \tag{2}$$

Each team of inventor types has a corresponding total net output, $q_k - c_k$, while the net output for individuals working alone is $q_{\underline{k}}(x)$. Individual types are assigned in the amount $N_k^{x,\ell}$ to a specific team type k. For instance, in a team with two of type (x, ℓ), $N_k^{x,\ell} = 2$. For each individual type (x, ℓ) on team k, the total output is shared such that there is no output left on the table. This means that the total wages paid out to team members is equal to the total team output:

[9] Teams also choose a patent technology class to work in. Given the team members, this optimal class would immediately follow. Because this immediately follows from the team composition, I leave this out of the model. I will be more specific on this problem when mapping this to its empirical counterparts in Section 5.3.

$$\sum_{(x,\ell)\in k} N_k^{x,\ell} w_k^{x,\ell} = q_k - c_k \tag{3}$$

Having discussed the net team output, $q_k - c_k$, and a general condition on wages, I turn to the individual's problem of choosing her team.

The Individual's Problem

Individual i is an infinitesimal agent of type $(x, \ell) \in \mathbb{X} \times \mathbb{L}$ who derives systematic and idiosyncratic utility from working in team type k. The systematic component, $w_k^{x,\ell}$, is a result of market forces and is the same across all types (x, ℓ) within a given team k. The second component, $\epsilon_k^{x,\ell}(i)$, is an iid preference shock for working in team type k which is specific for a given individual i of type (x, ℓ). This idiosyncratic iid individual-by-team utility is drawn from an iid type-I extreme value. This shock represents heterogeneous and unobserved reasons for forming teams. The distribution of the shocks is not related to the systematic observable component of an agent's skills. These shocks follow a cumulative distribution function as in McFadden (1974):

$$F(\epsilon) = \exp\left(\exp\left(-\frac{\epsilon}{\phi}\right)\right)$$

Each inventor i of type (x, ℓ) has the option to either work by herself or join a team. If she decides to work by herself, her team is indexed by $\underline{k}(x, l)$. If she decides to work with someone else, she can be matched with $\tilde{k}(x, \ell)$ team types, indexed by the other team member's type. Denote $K(x, \ell) = 1 + \tilde{k}(x, \ell)$ as the number of team types an individual of type (x, ℓ) can join. Due to the upper bound on team size $\overline{T}$ and finite number of types in the economy, there is a discrete set of team types i can join. Individual i observes a set of shocks across the team types they can join as follows:

$$\boldsymbol{\epsilon}^{x,\ell}(i) = \{\epsilon_k^{x,\ell}(i) : k \in K(x, \ell)\}$$

For each i of type (x, ℓ) there is a return equation for joining each team k:

$$\pi_k^{x,\ell}(i) = \underbrace{w_k^{x,\ell}}_{\substack{\text{systematic} \\ \text{component}}} + \underbrace{\epsilon_k^{x,\ell}(i)}_{\substack{\text{idiosyncratic} \\ \text{component}}} \tag{4}$$

Each individual i of type (x, ℓ) chooses her team k to maximize her return, $\pi_k^{x,\ell}(i)$. This is part of the equilibrium that is discussed next.

3.2 Equilibrium

The equilibrium is a set of wages across teams, the mass of individual types assigned to teams, and the mass of teams, $\{w_k^{x,\ell}, m_k^{x,\ell}, m_k\}$. The wages emerge as a result of the trading game to clear the market for each type and in each team. The endogenous assignment of types to teams, $m_k^{x,\ell}$, emerges from this process. The mass of a given team, m_k, is the mass of assigned types to the team divided by the number of unique team members.

The equilibrium is characterized by optimization of each agent i and market clearing within each team and within each type. The resulting equilibrium will have a sharing rule within each team type and frequency of each observed team type. The counterpart to each equilibrium object in the model will be explored in greater detail in the quantitative section when I map the model to patent data. I focus on the mass of each team m_k, the net return to each team $q_k - c_k$, and the expected value of being a given type (Proposition 3). Tracking these objects helps evaluate the changing patterns across teams and the effects of subsidies and expertise shocks on the economy.

Agent i observes her vector of idiosyncratic shocks and the systematic return to working for each team. She then chooses the team k that delivers the maximum return. Wages are determined endogenously by market clearing in teams for all types. There are five equilibrium conditions.

Definition of Equilibrium

- *Optimization:* Each $i \in \mathbb{X} \times \mathbb{L}$ chooses the team k^* to maximize the sum of her idiosyncratic and systematic income:

$$k^*(i) = \arg\max\{\pi_k^{x,\ell}(i) : k \in K(x,\ell)\}$$

This maximization delivers a relationship between wages and allocations that is governed by the dispersion of the preference shock ϕ and the mass of a given type (x,ℓ), $M_{x,\ell}$:

$$m_k^{x,\ell} = M_{x,\ell} \frac{\exp(w_k^{x,\ell}/\phi)}{\sum_{\tilde{k}\in K(x,\ell)} \exp(w_{\tilde{k}}^{x,\ell}/\phi)} \tag{E1}$$

- *Sharing Rules:* Wages for each agent on the team add up to total net output for each team k:

$$\sum_{(x,\ell)\in k} N_k^{x,\ell} w_k^{x,\ell} = q_k - c_k \tag{E2}$$

- *Market Clearing and Symmetry:* Markets clear for each inventor type (x,ℓ) (E3), the mass of each type assigned to each team is equal to the mass of the team multiplied by the number of this type on the team (E4), and there are no teams with negative mass (E5):

$$\sum_{k\in K(x,\ell)} m_k^{x,\ell} = M_{x,\ell} \qquad \forall \quad (x,\ell) \tag{E3}$$

$$m_k^{x,\ell} = N_k^{x,\ell} m_k \qquad \forall \quad (x,\ell) \in k \tag{E4}$$

$$m_k^{x,\ell} \geq 0 \qquad \forall \quad (x,\ell)\ ,\ k \tag{E5}$$

Proposition 1 follows from the equilibrium conditions.

Proposition 1. *An equilibrium that satisfies (E1)-(E5) delivers a relationship between the (i) masses of each team type k to (ii) the idea production function, (iii) the communication costs, and (iv) the supply of types as follows:*

$$\overbrace{\sum_{(x,\ell)\in k} \log\left(\frac{N^{x,\ell} m_k}{m_{\underline{k}}^{x,\ell}}\right)}^{(i)} = \frac{\overbrace{q_k - \sum_{(x,\ell)\in k} q_{\underline{k}}(x)}^{(ii)} - \overbrace{c_k}^{(iii)}}{\phi} \qquad s.t. \qquad \overbrace{\sum_{k\in K(x,\ell)} N^{x,\ell} m_k = M_{x.\ell}}^{(iv)}$$

Proof. See Appendix A.1. □

Proposition 1 links inventors' concentration in teams to the values of those teams; in particular, it relates the mass of a given team, m_k, relative to each type working alone to the value of the team q_k relative to working alone. Team frequency moves in a log-linear way with team output. As each agent becomes more productive alone (ii), or the communication costs increase (iii), the team will form less relative to each agent working alone. This framework provides the launching point of this paper as it will be used to quantify the forces behind the rise of teams and enable counterfactual studies, as (ii), (iii), and (iv) are linked to the benefits of teams, costs of teams, and supply of talent respectively. Next, I turn to two key elements of this framework that allow for identification and counterfactual analysis.

The local identification condition allows for estimation of the various forces driving the changes in the distribution of teams in the economy – benefits, costs, and supply. The counterfactual condition is amenable to discussing policies such as immigration that change the distribution of individual types in the economy.

Proposition 2. *Take a market with $\bar{T} = 2$ and the equilibrium from Proposition 1. Let* $\mathbf{M}$ *and* $\mathbf{m}$ *be the vector of the supply of each inventor type and vector of team masses, respectively. For* $\mathbf{M}^*$ *close to* $\mathbf{M}$, $\boldsymbol{m}^*$ *is uniquely determined.*

Proof. See Appendix A.1. □

This proof indicates that the equilibria are isolated. Beyond this analytical proof, Appendix A.1 shows the conditions under which this identification condition holds for teams of any size up to $\overline{T}$, which are met in the quantitative exercises in this paper. A key component of this proof is that the Jacobian determining the impact of supply shocks across types must be invertible. The Jacobian is the matrix of team masses thus its invertibility can be directly observed. This invertibility condition provides support for the counterfactual analysis in Section 6.

I now fix ideas of what the model attempts to capture in terms of quantifying the rise of teams. There are three pieces of the model that shift from the earlier period (e.g., 1980s) to the later period (e.g., 1990s). First, the idea production function $q_k - \sum_{(x,\ell)\in k} q_{\underline{k}}(x)$ may shift. Combinations of certain team types may yield different returns over time. At times, hardware producers see higher returns to working alone versus pairing together, or pairing up with a chemist.

Second, individual team types may face changing communication costs over time, c_k. For instance, with the advent of email and file sharing, inventors in separate locations can more easily produce together. This ability could drive down the costs of forming certain teams over time.

Third, the composition of inventors ($M_{x,\ell}$) may change. In particular, the 1990s saw a large movement toward fields related to computing, information storage, and hardware. This was driven by two forces beyond returns and costs. First, inventors selected into advanced degrees in these fields. Second, immigrant talent arrived with expertise in those fields, impacting the composition of inventors. Reductions in immigration restrictions in the US and corresponding outflows from Russia and China generated a large influx. As a result, the inventor composition in 2000 differed from the inventor composition in 1980. In Section 6, I discuss the influence of this migration and the role of self-selection. The following proposition will be an important reference.

The last result of this section characterizes the expected value of being an inventor of type (x, ℓ), which serves as an important statistic for informing immigration patterns and self-selection. This statistic for valuing types delivers a simultaneously intuitive and powerful predictor of the types that are the most productive in a large economy of teams. Proposition 3 discusses the properties of the ex ante value of being a specific type.

Proposition 3. *The expected value for an agent of type (x, ℓ) before her preference draws is as follows:*

$$\mathbb{E}[V_{x,\ell}] = cons + \underbrace{q_{\underline{k}}(x)/\phi}_{\text{output alone}} + \underbrace{\log\left(\frac{M_{x,\ell}}{m_{\underline{k}}^{x,\ell}}\right)}_{\text{concentration in teams}} \tag{5}$$

Proof. See Appendix A.1. □

Types that produce significant impact patents alone have higher value ex ante. Further, conditional on productivity alone, types who are frequently in teams receive relatively higher returns in teams. The respective weight is adjusted by the shock dispersion ϕ. This result is important for evaluating how self-selection impacts inventor migration in Section 6.

Section 5 uses the model framework to quantify the contribution of three key forces on the changes in the matching pattern. These three forces, returns, costs, and supply, exemplified by $q_k - \sum_{(x,\ell)\in k} q_{\underline{k}}(x)$, c_k, and $M_{x,\ell}$, respectively, impact the allocation to teams, m_k. Section 6 applies this framework to policy counterfactuals with a particular focus on immigration policy. I first discuss the data and build a bridge from the empirical framework to the quantitative analysis.

4 Data and Measurement

4.1 Data Sources

USPTO Patent Data and Inventor Disambiguation Algorithm

Although this paper uses several distinct data sources, it primarily relies on USPTO patent data for patents granted from 1975 to 2010. According to the USPTO, "US patent applications must list the 'true and only' inventors." A patent p is characterized by a technology class s[10], a team of inventors who jointly produce the patent, and forward citations, proxy for patent value, at subsequent dates.[11] I adjust for patent truncation using IPC1 patent class and the date of application in order to make comparable citations across different classes and time periods, as is standard in the literature (Hall et al., 2001).

Li et al. (2014) provide a dataset for inventor identification that links the entire career of an inventor to her history. They use a Bayesian disambiguation algorithm that employs patent class, location of inventors, firms where an inventor works, and her corresponding co-authors to track the full history of individuals on patents. The two main problems these algorithms deal with are cases of misreportings (e.g., misspellings such as "Jonh Smith") and common names (e.g., "John Smith"). The ability to identify inventors over time is crucial for building expertise measures that can speak to how inventors contribute to a team.

All types of assignees – firm, international, university, and government – are included. I trun-

[10]USPTO assigns a primary technology class in USPTO and WIPO assigns a primary IPC classification

[11]Patent stock market value and renewals are also used as measurement of patent value in robustness checks.

cate the data on both ends to capture experience vectors of individuals and clean citation data. The quantitative and empirical analyses focus on years 1980-2000, but inventor expertise vectors are based on the entire sample. The resulting sample includes 2.2 million unique patents, 1.5 million unique inventors – 1.1 million unique inventors who have more than 1 patent, and 4.5 million patent $\times$ inventor observations.

The class system and citation network admit identification of technological areas where inventors operate. For my measurement of skills, I take a stand on which level of classification to use to evaluate skill-sets. This paper explores three potential levels of classifications. The primary empirical analysis is done at the International Patent Classification 3-digit level (IPC3) for more granularity, but the quantitative analyses are done at the IPC2 level to have enough data for each type.[12] I also focus on USPTO classification measures in the empirical analysis robustness in Appendix C. The qualitative results do not shift significantly depending on the level of classification.

I use log patent citations as my measure of output quality, but test other measures in Section 7 and Appendix C. To adjust for the expected lifetime of citations, I use the truncated approach of Hall et al. (2001). I renormalize the value of sole-authored patents each period in order to have a similar benchmark, but I analyze other measures in Section 7 and Appendix C.

Records from the Soviet Union

To apply the model to immigration counterfactuals and verify its qualitative results, I make use of a unique dataset from the former Soviet Union. This dataset contains patent records from all inventions from the Soviet Union from 1924-1993. These documents are provided by Rospatent (Russian Federal Institute of Intellectual Property) and the Federal Institute of Industrial Property (FIPS).

FIPS provides all the data in the form of DVDs that contain complete scans of patent documents granted in the Soviet Union. Because the Soviet Union was essentially uninvolved in global technology production, this dataset provides a unique insight into the technologies produced in the Soviet Union. The DVDs record a total of 1.4 million unique documents.

There were two main types of patent documents in the Soviet Union: a patent and a certificate of authorship. A certificate of authorship was the most common patent document. These documents

[12] Appendix B.3 provides an example of the different levels of classification.

did not give an inventor the exclusive right to an invention, but the government did award prizes for inventions. Over the time period of the sample, there were a significant number of inventions, indicating that some incentives remained for inventors to innovate. My main goal with this data is to understand the domain of expertise of Soviet inventors at the time of the fall of the Soviet Union. As a result, I use the most recent innovation information (after 1970).

For the shock to be treated as quasi-experimental, it is necessary that inventors in the Soviet Union did not choose to innovate technologically with the expectation that their expertise as inventors would be integrated into the US market. This seems sensible given the unexpected nature of the fall of the Soviet Union and the fact that specialized human capital takes significant time to build. The discussion incorporates some of these forces. This is discussed further in Section 6, which addresses self-selection in migration.

Auxiliary Merged Patent Data

In addition to the standard patent measures and the disambiguation algorithm, I merge other datasets in order to speak to both robustness and ethnicity name matches.

Patent Value First, to ensure that patent citations are picking up similar outcomes to patent value, I merge in patent stock market value measures from Kogan et al. (2017). This measure delivers the projected value of a patent based on the change in stock market value of the day the patent was granted. This data is available on a limited set of patents, since it requires the firm be publicly listed. Nevertheless, it serves as a good verification exercise for observing the response of patent value to depth and breadth.

Ethnic Origin To match the ethnicities of inventors on patents for the purposes of both tracing out the ethnic diversity of inventors and matching ethnic Russians, I adopt the probabilistic ethnicity matching algorithm from Kerr (2007). This algorithm exploits the fact that certain names are common for certain ethnicities (e.g., “Wu” as Chinese or “Rodriguez” as Hispanic). The match rate for each name to an ethnicity is 80%. Next, I turn to the measurement of inventor and team skills, which will be important for the empirical analysis and quantitative results.

4.2 Data Measurement

This section constructs the empirical measures that are applied in the empirical and quantitative analysis. I start by building the measures of individual and team expertise and then proceed to discuss how measures at the individual-level contribute to measures at the team-level.

Inventor Expertise

Inventors hold skills that they deploy in the production of patents in teams. Because individuals work in teams, it can be challenging to extract their domain of expertise from team production. For my skill measure, I take the most prominent class of an individual, adjusting by team size for characterizing skill in class s as follows,

$$x_s = \sum_{p \in s} \frac{q_p}{T_p}.$$

Where q_p is patent impact (log lifetime citations) and T_p is the size of the team on the patent. To ensure the domain of expertise measure is robust, I build inventor types *adjusting by team size*. This down-weights skill measures in large teams. Additionally, I explore inventor types only using data from their *sole-authored patents* in Section 7. The tradeoff of counting skills in teams is losing information on individuals versus ensuring their skill is correctly identified. Fortunately, both sole-authored patenting and overall patenting deliver the same general result on the idea production function.[13]

- **Definition 1** *inventor skill in class s*

  Skill in technology s is given by an individual's total productivity in class s, x_s, net the focal patent. Where we recall $x_s = \sum_{p \in s} \frac{q_p}{T_p}$.

- **Definition 2** *inventor skill type x*

  Inventor i is skill type x if $\arg\max(x_1, x_2, ..., x_S) = x$

I distinguish between the total skill across technologies within an individual, which is multidimensional (Definition 1), with skill type x, which identifies the individual type and is relevant in

[13] I explore other measures in Section 7 and Appendix C. I discuss the counts in Appendix B.4.

Section 5.1. I apply an individual's most common technology class as her unique skill.[14]

Team Types

When individuals work in teams, they combine their skills. For instance, if one individual has skill x and another has skill y, they can form a team $k = (x, y)$ to produce output q_k. Individuals have a single skill type, but the type may contain a broad vector of skills.

This logic follows a similar structure to the model. A team is a vector of individual types across skills (e.g., technology class expertise). Given that teams form either to take advantage of deep or broad knowledge, I direct attention to team *depth* and team *breadth*, which measures the team skill on the focal patent.

- **Definition 3** *Team Depth*

  For a given team in technology x, team depth is the amount of expertise in x. Depth connects the team's expertise to the focal patent technology.

- **Definition 4** *Team Breadth*

  Team breadth is defined as the amount of skill outside the focal technology of a given patent. For a given technology x, I take their accumulated skill of types outside the focal technology.

Different teams may establish depth and breadth of expertise. Further, the changing nature of teams may generate changing returns to both of these forces. While teams form to take advantage of specific skills, the summary of team breadth and depth will help express the underlying causes of the rise in teams and specify an idea production function. The analysis in the following sections utilizes full information on individual expertise across classes but collapses this information into a specific type according to Definition 2. Analysis is done at the IPC 2-digit level.

Communication Costs

Individuals are identified by their skill type and location. I measure communication costs in two ways. First, as a discrete measure, I take the unique number of regions (4 regions within the US:

[14]To better understand the mechanism, Section 5.3 leverages an individual's bundle of skills as inputs into the idea production function.

East, West, Midwest, South, and one international, with 5 regions overall) on a patent. Second, as a robustness, I take the log of the expected geographical distance between skill types (e.g., the expected miles in between a chemist and biologist, see Section 7). Both measures deliver qualitatively similar results and indicate the rising communication between more distant inventors.

I focus on communication costs through geographical distance given the evidence that modern technologies shift primarily collaboration at a distance (Forman and van Zeebroeck, 2012). This paper presents a focus on geographical distance as we note this is the most core contribution of a fall in communication technology.

5 Quantitative Analysis

This section builds the quantitative and empirical framework that enables the study of the rise of teams and policy counterfactuals. To do so, I build a bridge from the matching model from Section 3 to empirical measures from Section 4. The goal of this section is two-fold. First, I provide the quantitative infrastructure to study the sources and implications of rising teams and the interaction of the skill distribution. Then, I quantify the contribution of each force (benefits, costs, inventor supply). Section 6 studies the implications of this framework in the context of immigration and taxation.

5.1 Quantitative Framework

This section maps the main three forces from the model (benefit, cost, supply) into data-relevant components. The following equation links the masses of each team type k (log count) – the outcome to (i) the *benefit* of teams, or the idea production function, (ii) communication *costs*, and (iii) the supply of types. There is also a residual component for each team k, ξ_k. Equation (6) summarizes the relationship between frequency of team and relative returns:

$$\overbrace{\sum_{(x,\ell)\in k} \log\left(\frac{N^{x,\ell} m_k}{m_{\underline{k}}^{x,\ell}}\right)}^{\text{log team count}} = \frac{\overbrace{q_k - \sum_{(x,\ell)\in k} q_{\underline{k}}(x)}^{\text{(i) benefits}} - \overbrace{c_k}^{\text{(ii) costs}}}{\phi} + \xi_k \qquad \text{s.t.} \quad \overbrace{\sum_{k\in K(x,\ell)} N^{x,\ell} m_k = M_{x.\ell}}^{\text{(iii) supply}}. \tag{6}$$

The mass of each team k is given by m_k, where the mass of type (x,ℓ) assigned to k is $N^{x,\ell} m_k$.[15] The relative benefit to teams in production comes from comparing the production in the team, q_k, to production alone for each type, $q_{\underline{k}}(x)$. As discussed earlier, the communication costs enter separately from the skill and are characterized non-parametrically by the number of regions team members live in.

Equations (Q1), which puts Equation (6) in regression form, and Equation (Q2) are the main equations for the quantitative analysis. These equations relate the estimated production function and communication costs across teams to the relative frequency of observing the team in the data (Q1), and estimate the underlying returns to teams on the skill-sets of inventors (Q2). I use the estimated production parameters that link the types of teams k to predicted relative output $\tilde{q}_{k,t}$ from Equation (Q2). I evaluate these two equations in the 1980s and 1990s:

$$\tilde{m}_{k,t} = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{q}_{k,t} + \tilde{\gamma}_t c_{n(k),t} + v_{k,t} \tag{Q1}$$

$$q_{p(k,s),t} = \alpha_{0,t} + \sum_{j=1}^{5} \alpha_{j,t} \mathbb{I}\{Quint(D_{k,s}) = j\} + \sum_{i=2}^{5} \zeta_{j,t} \mathbb{I}\{Quint(B_{k,s}) = j\} + u_p \tag{Q2}$$

I discuss Equation (Q1) and Equation (Q2) in turn. Equation (Q1) links estimated production function and the cost to the matching pattern. To be more specific, . $c_{n(k),t}$ is the number of regions covered by members of team type k. We classify types as discussed in Section 4.2 in order to estimate Equation (Q1). The components of Equation (Q1) are as follows:

$$\tilde{m}_{k,t} \equiv \sum_{(x,\ell)\in k} \log\left(\frac{N^{x,\ell} m_k}{m_{\underline{k}}^{x,\ell}}\right)$$

[15] This studies the rate of connection, e.g. how frequent is chemist-biologist both from the Midwest in the US.

$$\tilde{q}_{k,t} \equiv \hat{q}_{k,t} - \sum_{(x,\ell)\in k} \hat{q}^{x,\ell}_{\underline{k},t}$$

Equation (Q2) provides the estimation for the net benefit of team k, $\tilde{q}_{k,t}$. Equation (Q2) is estimated at the patent level, where I take quintiles of team depth and breadth. The coefficients on these respective quintiles, $\alpha_{j,s}$ and $\zeta_{j,s}$, are estimated within each IPC1 technology class. I fix the quintiles to match the entire time period of 1980-2000. This captures the fact that different technology classes exhibit different returns to team skills. Equations (Q1) and (Q2) are estimated over two periods, the 1980s and the 1990s.

Model Match

Individuals are identified by their type, which includes a bundle of skills, and location (x, ℓ). The mass of individuals of each type is given by $M_{x,\ell}$. For the quantitative decomposition, I split the sample into two distinct periods and rerun the model on those different time periods where individuals are identified through their sole-authored patents, delivering two distributions of individuals.

Individuals have one type of skill type, which is their primary technology background (e.g. where they have exhibited greatest impact). In order to group people into types, I take the average bundle of skills per type. The focus on tracking only sole-authored productions ensures the productivity is in the relevant arena. Additionally, individuals can only have one location which comes from where they primarily file patents (home address).

5.2 Quantitative Results

Before delving into the results, I discuss the intuition of how the quantitative framework elucidates the mechanisms discussed in this paper. We use the quantitative results to evaluate the three forces that could be changing the nature of teams, *benefits*, *costs*, and *inventor supply*. By shutting down each channel separately, we can ask what the counterfactual team size distribution is while only allowing the remaining two to change. Thus, each contribution (of %) to the change in team size isolates the contribution of the specific force.

Changes in the *benefits to teams* come from changes in the idea production function, $q_k - \sum_{x\in k} q_{\underline{k}}(x)$ for different types x. As certain teams generate higher quality patents, individuals will form in those teams more often. The strength of the response is governed by the dispersion of the preference shock ϕ. In Equation (Q1), $\tilde{\beta}_1$ governs this responsiveness. There are two main sources within the idea production function that could be driving the changes in team size through the idea production function. Teams may exhibit larger returns to depth – ideas require deeper expertise and so inventors choose to form teams with those somewhat similar to themselves. Second, teams may exhibit larger returns to breadth – ideas require more expertise from different patent classes so inventors work with those not in the same class. I find both of these forces contribute to the changing idea production function. Yet, as I show in Section 5.3, the returns to *depth* of expertise is the most relevant changing component at the end of the 20th century.

Team formation also responds to the geographical distance between its members, $d_{k,t}$. $\tilde{\gamma}_t$ in Equation (Q1) governs the responsiveness of team frequency to communication costs, adjusted by team size T_k and the dispersion of the idiosyncratic shock, ϕ. If all inventors in the 1990s had been distributed across the map in the same way as they were in the 1980s, then a fall in $\tilde{\gamma}$ would have generated larger teams at greater geographical distance. The overall quantitative goal is to ask about the contribution of these forces to changes in the matching pattern.

Using the machinery from the empirical analysis, I proceed by estimating the parameters of interest: $\alpha_{j,s,t}$, $\zeta_{j,s,t}$ come from the production function equation. $\tilde{\beta}_{1,t}$ and $\tilde{\gamma}_t$ come from the matching equation and identify ϕ and β_c in the underlying model.

Using log lifetime citations as the outcome variable of interest, I perform regressions at the IPC1 level to retrieve depth and breadth coefficients in a nonparametric fashion separately for the 1980s and 1990s across technologies. Teams have a given depth and breadth across each class s. Having illustrated the match between the estimated and realized production function, I turn to the quantitative properties of how production and cost are linked to matching from Equations (Q1) and (Q2).

In addition to the production function, the key additional variables of interest in the quantitative procedure are the coefficient on coordination costs β_c and the dispersion of the preference shocks ϕ. Figure 5 illustrates how the model matches the data. Here, I take the fitted values of the team from the previous equations and plot them against the realized team matches. The R^2 is 0.43.

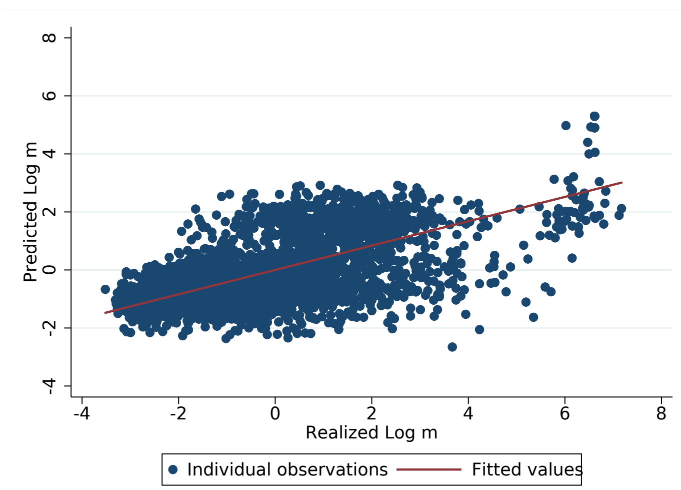
A point worth stressing from Figure 5 is that the predicted and realized matching has no mechanical relationship. This is because the model only delivers coefficients that govern the noise $\tilde{\beta}_1$

Table 1: Parameter Values

Parameter	Description	Value	Main Identification
		— Panel A. From Production Function —	
$\alpha_{j,s,t}$	Quartiles of depth in patent class s	$4 \times$ num. major tech $\times$ 2 periods	Production Estimation (Q2)
$\zeta_{j,s,t}$	Quartiles of breadth in patent class s	$4 \times$ num. major tech $\times$ 2 periods	Production Estimation (Q2)
		— Panel B. From Matching Function —	
$\gamma c, n(t)$	Coefficient on communication cost	2×3 coefs	$\tilde{\beta}_{2,t}$ from Equation (Q1)
ϕ	Preference shock dispersion	0.49^{***}	$\tilde{\beta}_{1,t}$ from Equation (Q1)

Notes: This describes the estimation and description of parameter values in the quantitative analysis. Column (1) indicates the parameters from Equations (Q1) and (Q2). Column (2) describes these coefficients. Column (3) provides the value or number of parameters (if more than one). Column (4) indicates the method of identification.

Figure 5: Production, Cost, and Realized Frequency of Team Type



Notes: This figure plots the predicted and realized density of each team type in both periods in the quantitative model. Equation (Q1) lists the log relative count of each team (x-axis) and Equation (Q1) predicts the relative share of each team (y-axis). This relationship is untargeted beyond the parameters from Table 1. The R^2 is 0.47. Source: USPTO and author calculations.

and communication costs. The positive correlation comes from linking the expected returns and the matching pattern. Overall, teams with higher expected returns are more likely to match. This is unsurprising, but does not directly follow from any mechanical assumptions.

Results: Idea production function, communication costs, and supply of skill.

Once I estimate the parameters governing the production and matching equations, I shut down each channel to understand its implied contribution to the change in team size. I start with a counterfactual analysis in order to understand the contribution of each force to the change in team size. I run the model for the two periods separately in order to uncover the changing parameters. This is enabled by the parameters of the two main identifying equations, Equation (Q1) and Equation (Q2).

I ask how much would the size of teams change by shutting down each channel on its own. In the case of the benefits of teams, I keep the parameters from the production function pre-period ($\alpha_{j,s,t}, \zeta_{j,s,t}$) and compare the realized to the counterfactual team size. I do the same thing with the coefficient on communication costs ($\tilde{\gamma}_t$) and the shares of types ($M_{x,\ell,t}$) respectively. Table 2 shows how much of the change in team size these three forces can explain.

Table 2: Contribution of each force to changes in matching pattern

Causal Force	Share of Change in Explained
Benefit (Idea Production Function)	68%
Cost (Communication across regions)	25%
Composition	-9%
Unexplained	16%

Notes: This table runs the model as described in the main text and asks how much each force contributes to the observed share in team size. Source: USPTO and author calculations.

Changes in the idea production function can explain 68% of the change in team size from the 1980s to the 1990s. Falling communication costs, as understood through regional dispersion, explains 25% of the change in team size. Composition pushes in the other direction in the quantitative exercise, driving -9% of the change according to IPC2 classifications.

This quantitative result finds that changes in the idea production function are the most salient force behind the changes in team size. It illustrates that while communication costs are indeed falling, there is an increasing premium on understanding policies that foster complementary skills in teams. I turn to some reduced form empirical analysis to confirm and express the main mechanism driving shifts in the benefits and costs of teams.

5.3 Testing Quantitative Mechanism

The quantitative analysis indicated that the rising benefits to teams is the most prevalent force driving the change in team size, yet communication costs also played an important role. The decomposition allows us to quantify each role, but we did not directly focus on how production changed.

This section focuses on some reduced form empirical results to test the overall quantitative mechanism. In Appendix C, I focus on the *costs* of forming teams and the compositional effects in a reduced form way. I start by characterizing the returns to depth and breadth in the 1980s and the 1990s, to focus on the impact of patents depending on the skill-set of individuals on the team.

Empirical Result: Teams with higher depth and breadth produce more impactful patents, and even more so in the 1990s than in the 1980s.

To evaluate the regression with depth and breadth as the driving force behind production, I take log citations[16] as the y-variable, exploring other variables in Appendix C. As inputs, I take the depth and breadth at the patent-team level and build deciles of each over time. Equation (7) illustrates the specification:

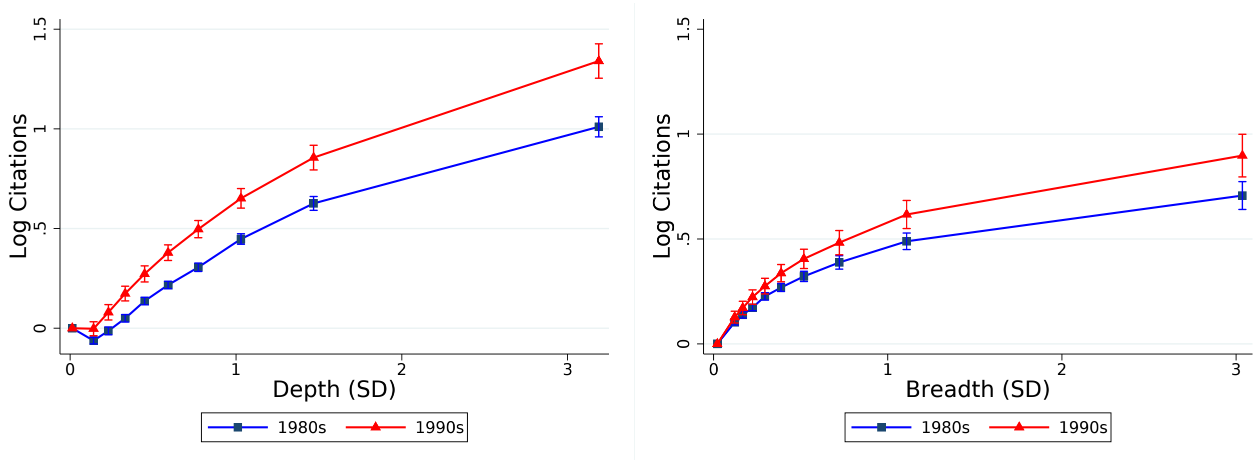
$$y_{p(k,s)} = \alpha_0 + \sum_{j=2}^{10} \alpha_j \mathbb{I}\{Decile(D_{k,s}) = j\} \times 1990s + \sum_{i=1}^{10} \zeta_j \mathbb{I}\{Decile(B_{k,s}) = j\} \times 1990s + \mathbf{Z}_{p(k,s)} + u_p \tag{7}$$

Equation (7) includes an interaction term for the 1990s. The goal is to outline how log citations respond to depth and breadth across the two different time periods. If there were a change to the idea production function over time, then one would expect higher coefficients on depth and breadth in the 1990s relative to the 1980s.

Indeed, as indicated in Figure 6, higher-decile depth and breadth are linked to higher-quality patents in the later period. These results illustrate that the returns to skill are changing over time, as

[16] To be consistent with the literature and account for zeroes, I take the inverse hyperbolic sine transform.

Figure 6: Citations and Depth/Breadth, 1980s versus 1990s



(a) Depth

(b) Breadth

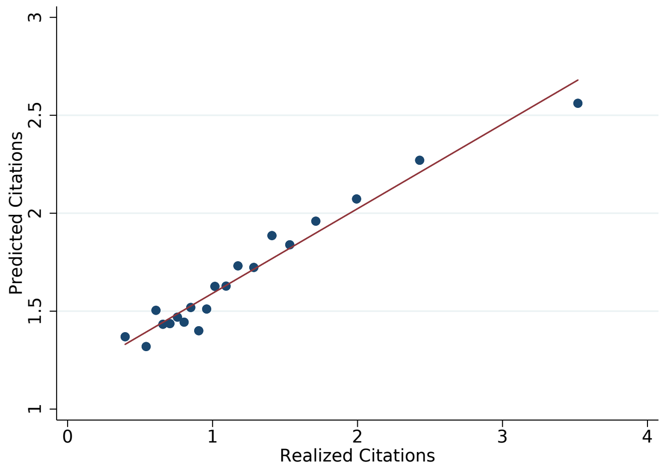
Notes: This figure plots the coefficients from Equation (7) split by time period. Skills are accumulated from all time periods for analysis, but regression only includes lifetime citations from 1980-2000. Source: USPTO and author calculations.

patents exhibit a stronger response to team depth and breadth. This is true even when controlling for team size, technology, and year effects. It also holds when building quintiles by period. These results indicate that the changing nature of the benefits of teams come from the skills individuals combine.

One additional point to note is that the predicted value of patents from this simple breadth depth model is a good fit overall for predicting patent quality, which was applied in Section 5.2. Figure 7 plots the estimated production value of teams (y-axis) against the realized value for those teams observed in the data. This estimation procedure allows me to input predicted values for noisy teams and teams not realized in the data in the pre- or post-period.

This section characterizes the gross output measures for various collections of skills of teams, q_k. I show in Appendix C that communication costs have negligible effects on patent quality, so the main focus on the benefits of teams is through skill, though I leave a discussion to endogenous location sorting for later research. To preview the interaction of migration, policy, and innovation, I turn to policy applications next.

Figure 7: Realized and Model-implied Citations



Notes: This binned plot contains about 500 observations per point which represents the average citations per team. Source: USPTO and author calculations.

6 Policy Applications

The previous section built a framework to quantify the sources behind the rise of teams. Further, I illustrated that the rise of teams is connected in a significant manner to changes in the idea production function. This implies team production is an essential input into innovation policy analysis. The framework in this paper is amenable to a wide variety of policy analyses. What does this imply for economic policy? How does taking teams and matching seriously affect classical policy issues, such as taxation and immigration?

I focus on taxes in Section 6.1. I then turn to high-skilled immigration, making use of the fall of the Soviet Union as a historical episode in Section 6.2. In Appendix D, I discuss more general lessons for this framework for immigration policy and the implications for R&D and education policies.

6.1 Tax Policy

To understand the effects of taxes on innovative output, I take the later period distribution of skills and team composition and ask how taxes affect team composition. One advantage of this framework is it allows researchers to isolate the *sorting* channel in understanding how taxation affects innovation. The counterfactual sorting from taxation comes from the equalization of returns across teams (e.g., 100% taxation does not incentivize inventors to find the "best" team, in the output sense).

Tax policy can be introduced in two ways. First, inventors can be taxed on their *gross income*, i.e. not include their costs of team formation if not able to be put in their budget. Second, inventors can be taxed on their *net income*, assuming communication costs are deductible in their taxes. We focus on a tax that hits wages initially, such that for team k skill x:

$$w'_{x,\ell,k} = (1-\tau) w_{x,\ell,k}$$

Because there is no endogenous aggregate labor supply in this model, the shifts in taxes shift the distribution of teams. When individuals choose teams, they now weigh the overall return to each team. Large taxes induce more "random" team matching as they induce individuals to be more indifferent between teams. For instance, a 100% taxation leads to a fully random distribution of teams in proportion to the supply of skill in the economy.

I return to the matching equation to explore the effects of these taxes on wages paid out. This taxes the *net* return to each team, which induces the following matching pattern:

$$\sum_{(x,\ell)\in k} \log\left(\frac{N^{x,\ell} m_k}{m_{\underline{k}}^{x,\ell}}\right) = \frac{(1-\tau)\left(q_k - \sum_{(x,\ell)\in k} q_{\underline{k}}(x) - c_k\right)}{\phi} \qquad \text{s.t.} \qquad \sum_{k\in K(x,\ell)} N^{x,\ell} m_k = M_{x.\ell},$$

Where aggregate innovation is $Q = \sum_{k\in\mathcal{T}} m_k q_k$. Qualitatively, we can see that tax policy attenuates the sorting of individuals to teams. To see the quantitative results of this policy, I input 10%, 20%, 50% taxes into the existing pattern of matches and output measures from the 1990s period. Table 3 shows the allocation of individuals to teams based on the taxes. It is not surprising that the taxes induce lower innovation. Yet a key element to note is that this tax is a wedge on the matching pattern and does not incorporate endogenous labor supply, which is a key driver of the previous papers on taxation and innovation empirically and theoretically (Akcigit et al., 2016; Jaimovich and Rebelo, 2017). Further, this allows for more detailed counterfactual analysis of tax policy which is of rising interest amongst economists studying innovation (Akcigit and Stantcheva, 2020; Akcigit et al., 2021).

Table 3: Taxes and Aggregate Innovation

Tax rate	Aggregate innovation	% change from baseline
0%	1	0
10%	0.93	-7%
20%	0.86	-14%
50%	0.64	-36%

Notes: This table compares the aggregate innovation from the matching pattern and idea production with different levels of taxes. Rows 2-4 introduce taxes into the matching model and ask about the counterfactual allocation and corresponding innovative output. Source: USPTO, FIPS, and author calculations.

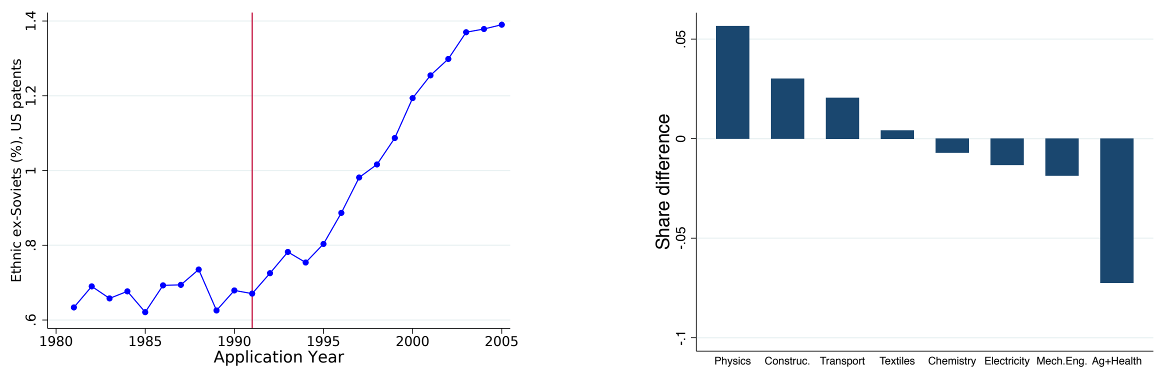
Table 3 expresses a simple result. The changes in tax policy indicate that to understand the effects of taxes policymakers should think through the sorting of individuals to teams. This is an element that must be kept in mind for policymakers interested in raising revenue through taxing inventors.

There are many possible additions to this policy. It is possible individuals have interest in sorting to higher productivity teams for reasons other than the private return. This would attenuate the result, and could be compared internationally in terms of optimal composition across tax regimes. Going in the other direction, high enough taxes may induce labor supply reductions of inventors. Both avenues will be interesting to observe into the future. I now turn to understand a real-world immigration shock to learn more general lessons about migration and innovation.

6.2 Real-World Immigration Shock

Inventors often move across borders and immigration policies can play a significant impact on the global distribution of talent (Akcigit et al., 2016; Kerr, 2018). Recent studies (e.g. such as, Burchardi et al., 2020) find significant effects of immigrants on innovation. This section applies a key point in this paper to build a bridge to this literature. To understand the contribution of immigrants, it is essential to understand that immigrant inventors work in teams. To understand immigration through the lens of specific expertise and team production, I make use of a historical episode and new data from the Soviet Union discussed in Section 4. After the fall of the Soviet Union, there was a large influx of Russian inventors into the US, as can be seen in Figure 8. Figure 8a plots the proportion of Russians on US patents in the 1980s and the sudden uptick post-1991 when the Soviet Union fell. Figure 8b shows the differential expertise of the US and the Soviet Union in the 1980s. This figure provides a promising example of a talent supply shock (as seen in Figure 8a) of different types of inventors (in Soviet Union versus US as in Figure 8b).

Figure 8: Fall of the Soviet Union



(a) Ethnic Russians on US patents

(b) Tech. Diff. w/ US (Soviet Records)

Notes: Panel (a) plots the share of ex-Soviet bloc ethnicities on US patents in the USPTO. Panel (b) shows the technological distribution of patents in the Soviet Union compared to patents in the US in the 1980s by IPC1 technology category. Source: USPTO, FIPS and Kerr (2007).

I use this shock to validate elements of the model and provide suggestive implications for immigration policies, in a similar manner to previous work studying the impact of Russian mathematicians on the west (Borjas and Doran, 2012; Agrawal et al., 2016). I start by shocking the distribution of expertise in the US economy, proportion to the amount of newly arrived Russians from 1995-2005 (as seen in Figure 8a), in the 1990s with a sample representative of the Soviet Union (as seen in Figure 8b). This delivers the projected aggregate innovation of an increase in the supply of Russians by 0.7% of the US population which captures their expertise from the Soviet Union.[17]

I compare the predicted and realized contribution of immigrating Russians using patenting by newly arrived ethnic Russians in USPTO data. In this exercise, I remove the existing set of

[17] I abstract away from the spatial distribution of inventors for this exercise and focus only on the expertise.

Russians with their expertise and project the counterfactual aggregate innovation. For both these exercises, it is crucial to define aggregate innovation, which is as follows:

$$Q = \sum_{k \in \mathcal{T}} m_k q_k$$

Understanding matching of team types, m_k, and idea production, q_k, enables the study of different distributions of expertise in the economy and their contribution to aggregate innovation. After studying the response of the economy to the Russian influx, I turn to general principles that provide qualitative insights for economic policy.

To identify USPTO inventors who are Russian, I link ethnicity probabilities to inventors in the US using a procedure from Kerr (2007). I classify an inventor as a certain ethnicity if their name delivers a greater than 0.5 probability of the given ethnicity through first and last name match. In addressing how Russian expertise contributed to US innovation, I compare the distribution of the Soviet Union across IPC3 patent classes to the US distribution across IPC3 classes in the 1980s. I perform the same comparison with ethnic Russians in the US from 1995-2005 whose first patent was produced past 1991. There is no way to match specific names from the Soviet Union to the US given the common name changes that took place as Russians moved. To specify the probable immigrants, I use the ethnicity of "new" patenters of Russian ethnicity to infer whether they are from the Soviet Union.

As Borjas (1987) notes, immigrants to the United States are self-selected. I model this through a moving cost ψ which each immigrant faces. This cost will induce selection for immigrants associated with skills that have more value in the US. Since Proposition 3 delivered a rank order of the value of expertise, I leverage this to generate the private value in the US of being a specific type of inventor. Those with expertise in low value patent classes may not find it worth it to move given the small change in returns by moving to the US. Equation (8) returns to this result:

$$\mathbb{E}[V_x] = cons + \underbrace{q_{\underline{k}}(x)/\phi}_{\text{value alone}} + \underbrace{\log\left(\frac{M_x}{m^x_{\underline{k}}}\right)}_{\text{value from teams}} - \underbrace{\psi}_{\text{moving cost}} \tag{8}$$

Due to the moving cost being unobserved, I explore different implied ψ cutoffs. When presenting the impact of Russian migrants, I focus on different implicit moving costs that would tend to draw high types as in Equation (8). Due to the flat returns to innovation in post Soviet Russia, it is sensible that those with higher value technologies in the US would be more likely to migrate. I confirm this by asking how correlated the predicted distribution of Russians across patent classes is to the realized distribution.

Table 4 compares the two exercises and their correlation. Panel A shows the contribution of Russians to aggregate innovation predicted by the Soviet Union records with varying degrees of self-selection. The third column compares the predicted distribution of experts to the realized distribution. For the predicted change using Soviet data, I show how the output response changes depending on the degree of self-selection (no self-selection, and the top 50 and 20 IPC3 classes as cutoffs respectively). Panel B delivers the model-implied realized contribution of newly arrived Russians in teams. Note this value is closer to the predicted value with significant self-selection.

Table 4: Contribution to Aggregate Innovation, 1995-2005

Measure	**Δ Agg. Innov (%)**	***Corr(Pred. SU, RU in US)***
	— Panel A. Predicted Impact, Soviet Union shock —	
No selection	0.55	0.69
Selection at T50	0.77	0.86
Selection at T20	1.08	0.88
	— Panel B. Impact of Soviet Union shock —	
Russian inflow	0.81	1

Notes: Calibrated model output using counterfactual skill distribution. In Panel A, each row represents a simulated distribution of teams adding the shock from the Soviet talent, with different selection thresholds (e.g. proportional to selection) from Equation (8). In Panel B, the actual Russian distribution is simulated. Source: USPTO, FIPS and author calculations.

I stress two main results from Table 4. First, the Russian contribution to aggregate innovation was 15% greater than their increase in the population in terms of inventors (0.81 vs. 0.70). Second, this was due to the self-selection of Russian migrants, which is a key element of migration generally; individuals migrate depending on the available teams and impact of producing alone (almost 90% of Russians produced in teams). If the US distribution were shocked with the same mass of new inventors distributed as in the Soviet Union, it would understate the Russian contribution by over 50% (0.55 vs. 0.81). This result is robust to different skill measures and output measures, as discussed in Section 7 and Appendix D.1. To understand the role of immigration in this setting, researchers and policymakers must take into consideration the distribution in the Soviet Union, the degree of self-selection, and the team contribution channel. This paper provides a quantitative framework to do so.

This result contains more general lessons for policymakers who want to leverage immigration to increase innovation. First, a general increase in high-skilled immigration will tend to contribute more than purely predicted through the skill distributions due to self-selection. Second, for targeted policies, there are simple statistics that can help policymakers understand the skills most in demand. If policymakers want to target specific skills, they can leverage Proposition 3 and Equation (8).[18] Policymakers with incomplete information about the structure of the innovation economy can rely on two straightforward pieces of information from Equation (8): how productive is a specific inventor type (e.g. organic chemist) alone, and how often are they in teams?

7 Discussion and Robustness

The quantitative results developed in Section 5 and Section 6 deliver clear implications of the contribution of various forces to the rise and teams and a framework to analyze tax and immigration policy. The results are based on the empirical measures discussed in Section 4. To ensure robustness to different aspects of the data, this section focuses on varying measures of inputs and outputs, changing the time window, and the geographical distance measure. Further, I evaluate in greater detail concerns about measurement and endogeneity.

First, I ask how the outcome changes with a different measure of patent quality–the private value of patents from the stock market and the *level* of citations. Second, I ask how measuring individual skills only including sole-authored patents (so as to not introduce any possible skill mismeasurement due to team complementarities) affects the outcome measure. Third, I focus on measuring the average distance across types by distance (e.g. average log miles across types) as a measurement of the communication cost. Fourth, I focus on extending the time period to see if we can garner more general lessons into the 21st century. Lastly, I focus on potential concerns around endogeneity of the skill measure, and the cases in which it would be concerning, and show that the sources of endogeneity do not lead to concerns.

[18] Appendix D discusses these issues in greater detail.

Overall, the main messages from the previous sections are robust to these different measures, though the quantitative results are slightly different. The changing nature of the benefits of teams is the largest driver behind the rise of teams; taxes have significant effects on production simply through the sorting channel; Russian immigrants produce greater innovation than expected due to self-selection and team formation. This can be seen in Table 6, which summarizes the main differences depending on changes in the measures.

Measure of Patent Quality

One question may emerge that the patent quality measure (e.g., log citations) may not measure quality relevant to agents on the patent. I address these concerns in a few ways. First, I take the stock market value of a patent, applying a measure from Kogan et al. (2017) . This measure is invariant to time period, so one does not need to worry about renormalization. Further, since this captures the private value of patents, it should affect the sorting behavior of teams. Second, I look at the *level* of citations, re-normalized each period. This can be seen in rows (ii) and (iii) in Table 6 respectively. I find similar results in both cases when it comes to the forces behind the rise of teams, and implications in tax and immigration policy.

Measure of Skill by Type

As discussed in this section, to measure individual skill type I take the most prominent class of an individual, adjusting by team size for characterizing skill in class s as follows,

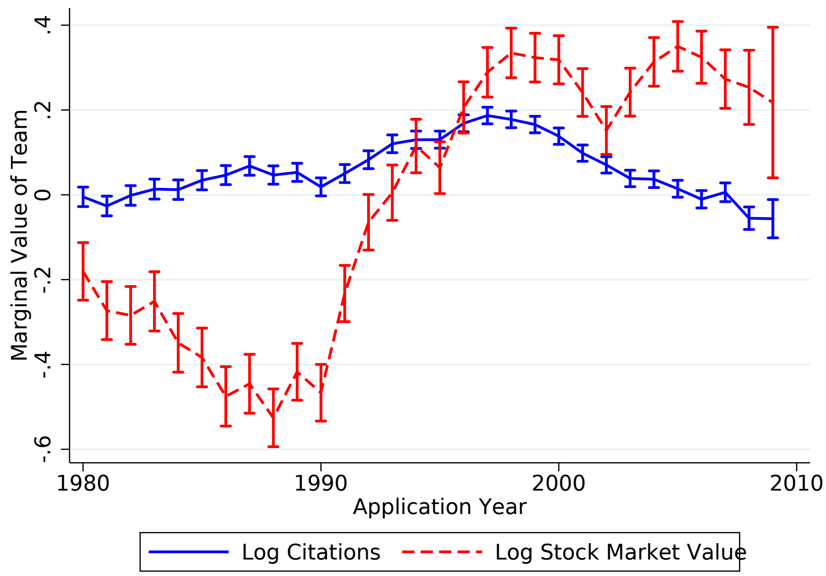
$$x_s = \sum_{p \in s} \frac{q_p}{T_p}.$$

Where q_p is patent impact and T_p is the size of the team on the patent. For the skill in each class by type, I take the average skill vector within each type. Yet, there may be concerns about extracting an incorrect skill through the team collaboration. I turn attention to sole-authored patents only in order to ensure similar results when we have a correct definition of skill. The results are qualitatively similar, as we see in Table 6.

Time Periods

This paper compares the 1990s to the 1980s to deal with truncation issues. However, the rise of team size is part of a long-run trend. From 1980-2010, the average team size increased, as citations across teams against individuals increased and communication costs fell. Figure 9 shows this trend, and shows persistence in both stock market value and log citations into the 2000s.

Figure 9: Relative Citations and Stock Market Value, Team to no Team



Notes: This figure plots the relative value of team to no-team in USPTO patents. The log lifetime citations suffer from truncation issues in the later decade. I follow Hall et al. (2001) in truncation. Source: USPTO and Kogan et al. (2017).

Endogenous Team Formation

The framework in this paper incorporates endogenous selection into teams. One concern may be related to the fact that individuals select into teams based on margins that are unidentified in the data (e.g., certain types within types). I address this on two margins. First, I look across the *vertical* dimension of skill, to see whether more skilled people working alone are more or less likely to produce higher quality innovations in teams. I find that across these measures of skill there is not a significant difference in the propensity to join teams, and it does not change across time periods. Second, I ask whether individuals working on teams of various vertical skills show heterogeneous returns to teams. Similarly, the effects do not appear to be significant. I address this in Equation (9):

$$y_{p(k,s)} = \alpha_0 + \sum_{j=2}^{3} \alpha_j \mathbb{I}\{Terc(S) = j\} \times \mathbb{I}\{team\} + \sum_{j=2}^{3} \alpha_j \mathbb{I}\{Terc(S) = j\} \times \mathbb{I}\{90s\} + \mathbf{Z}_{p(k,s)} + u \tag{9}$$

Where $y_{p(k,s)}$ is the outcome of interest, where is the indicator of whether the patent is a team patent or the quality of the patent. In the case where the team indicator is the outcome of interest, the team interaction on the right-hand side is dropped. Skill S is the total patent impact of the individual, net the focal patent. $\mathbf{Z}_{p(k,s)}$ represents patent-level controls for technology, type, and year. The coefficients from Equation (9) can be found in Table 5. We focus in particular on the interaction between the skill bins and the marginal effect on patent quality. In particular, the marginal returns to teams appear to be unrelated to the vertical dimension of skill.

Table 5 shows that the propensity to join teams conditional on skill level is not significantly different for different terciles of *vertical* skill, when technology is controlled for. However, it appears the returns to skill are marginally lower for the top 2 terciles of skill (about 3-7% lower for the top two terciles of skill) to joining teams than those not at teams. However, since the propensity does not differ, the interaction point of these forces does not appear to be an important ingredient in the changes in team size.

Robustness of Quantitative Results

Table 6 compares the benchmark calibration outcomes to their outcomes when a new measure is introduced. The table takes the (i) benchmark calibration and then performs the same analysis with (ii) the stock market value as a measure of quality; (iii) citation level (instead of log) as measure of quality; (iv); including only sole-authored production to measure skill, and (v) using average distance across types instead of unique count of regions.

With each of these adjustments, I re-evaluate the main three results of the paper and find they are fairly robust to the shifting definitions. First, I evaluate the joint contribution of benefits, costs, and inventor composition (supply) to the shifting size of teams in the last two decades of the 20th century. Second, I evaluate the innovation impact of taxation and compare the innovation cost of a

Table 5: Effect of Skill Type on Propensity to Join Team and Patent Quality

	(1) Join Team	(2) Log Citations
Skill Type 2	0.009	0.182***
	(0.005)	(0.018)
Skill Type 3	0.008	0.326***
	(0.009)	(0.034)
Skill Type 2 × 1990s	0.005	0.040***
	(0.004)	(0.01)
Skill Type 3 × 1990s	0.011	0.118**
	(0.006)	(0.039)
Skill Type 2 × team	.	-0.039**
	()	(0.012)
Skill Type 3 × team	.	-0.072**
	()	(0.025)
Observations	3338119	3338119
R^2	0.051	0.121
Technology/Year Controls	Y	Y

Notes: Column (1) treats the team as a y variables in Equation (9). Column (2) looks at citation output and interacts the skill tercile with team. Robust standard errors in parentheses.
*, **,***: Significant at 5%, 1% and 0.1% level respectively.

10% tax on net output to the benchmark. The fifth column studies the relative impact of Russians compared to their predicted impact, and I find that in all specifications the realized impact is larger than the predicted impact not taking into account self-selection and team production.

There are a few takeaways from Table 6. First, the changes in idea production seem to be the strongest force in shifts to team production, but in some cases communication costs appear more significant than others, yet never more significant than changing idea production.[19] Second,

[19] For the decomposition, there is still residual unexplained variation from specific teams that occur more or less frequently than the quantitative model predicts. As a result, the percentages in the three rows do not necessarily add up to 100%. Indeed, some rows on their own might predict a larger than 100% effect.

Table 6: Robustness – Decomposition and Policy Counterfactuals

	Benefits	**Costs**	**Supply**	**10% tax cost**	**RU impact**
(i) Benchmark Calibration	68%	25%	-9%	7%	1.56
(ii) Stock Market Value as Quality	102%	25%	-5%	7%	1.27
(iii) Citation Level as Quality	44%	25%	3%	2%	1.13
(iv) Sole-Author as Expertise	73%	25%	7%	15%	1.04
(v) Average geo distance	47%	36%	-2%	Same	Same

Notes: This table reruns the main analysis in the paper varying the definition of skill and output. Source: USPTO, FIPS and author calculations.

self-selection in immigration and taxation both have significant impact on aggregate innovation in teams. This indicates the sorting pattern has important policy implications that can't be ignored from an innovation perspective. Lastly, when I change the definition of geographical distance, it does not affect the tax or immigration policy, since these policies were analyzed with the location information removed.

8 Conclusion

Complex tasks in the economy increasingly require more varied skills and larger teams; this is particularly salient in the case of innovation. This paper addresses the forces that underlie the increasing importance of teams in innovation as well as their macroeconomic and policy implications by building a quantitative framework of matching and innovation in teams.

I proceed by building a team idea production and matching framework. With this framework and USPTO patent data, I quantify the role that three major forces play in driving these patterns: benefits (the idea production function), costs (communication costs), and supply (inventor expertise composition). I find that all three forces are relevant for both the technological composition of the economy and the prevailing team size; changes in the idea production function, as understood through the returns to teams, explain the largest change in team size.

Given the rising importance of fostering complementary skills, the results have relevant policy implications. The model provides immediate insights on both taxation and immigration policy. For taxation, the interaction of taxes with sorting into teams has first-order effects on aggregate innovation. For immigration, the model provides a good basis for policymakers interested in skills-based innovation policy and a method for modeling how self-selection of immigrants interacts with overall innovation.

This framework suggests further avenues to explore. For instance, while this model delivers a distribution of expected values across skills, the cost of training each of these skill types is important to know for questions of skill investment and education policy. Understanding the interaction of the cost of training skills and the innovative output is a fruitful area for research. Lastly, this paper provides a promising framework to understand the incentives of firms to collect teams and the interaction of firm dynamics with team dynamics. Thus, this paper can serve as a first step to quantify the role of teams in innovation and provide a framework for a broad range of investigations.

References

Adams, J. D., G. C. Black, J. R. Clemmons, and P. E. Stephan: 2005, 'Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999'. *Research Policy* **34**(3), 259 – 285. University-based Technology Initiatives.

Aghion, P. and P. Howitt: 1992, 'A Model of Growth Through Creative Destruction'. *Econometrica* **60**(3), 323–351.

Agrawal, A. and A. Goldfarb: 2008, 'Restructuring Research: Communication Costs and the Democratization of University Innovation'. *American Economic Review* **98**(4), 1578–90.

Agrawal, A., A. Goldfarb, and F. Teodoridis: 2016, 'Understanding the Changing Structure of Scientific Inquiry'. *American Economic Journal: Applied Economics* **8**(1), 100–128.

Ahmadpoor, M. and B. Jones: 2017, 'The Dual Frontier: Patentable Inventions and Prior Scientific Advance'. *Science* **357**(6351), 583—587.

Akcigit, U., S. Baslandze, and S. Stantcheva: 2016, 'Taxation and the International Mobility of Inventors'. *American Economic Review* **106**(10), 2930–81.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi: 2018, 'Dancing with the Stars: Innovation Through Interactions'. Working Paper 24466, National Bureau of Economic Research.

Akcigit, U., J. Grigsby, and T. Nicholas: 2017a, 'Immigration and the Rise of American Ingenuity'. *American Economic Review* **107**(5), 327–31.

Akcigit, U., J. Grigsby, and T. Nicholas: 2017b, 'The Rise of American Ingenuity: Innovation and Inventors of the Golden Age'. Technical report. National Bureau of Economic Research WP23047.

Akcigit, U., J. Grigsby, T. Nicholas, and S. Stantcheva: 2021, 'Taxation and Innovation in the Twentieth Century*'. *The Quarterly Journal of Economics* **137**(1), 329–385.

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

Akcigit, U., J. G. Pearce, and M. Prato: 2020, 'Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth'. Working Paper 27862, National Bureau of Economic Research.

Akcigit, U. and S. Stantcheva: 2020, 'Taxation and Innovation: What Do We Know?'. Working Paper 27109, National Bureau of Economic Research.

Azoulay, P., J. S. Graff Zivin, and J. Wang: 2010, 'Superstar Extinction'. *Quarterly Journal of Economics* **125**(2), 549–589.

Barro, R. J.: 2001, 'Human Capital and Growth'. *American Economic Review* **91**(2), 12–17.

Becker, G. and K. M. Murphy: 1992, 'The Division of Labor, Coordination Costs, and Knowledge'. *Quarterly Journal of Economics* **107**(4), 1137–1160.

Becker, G. S.: 1973, ‘A Theory of Marriage: Part I’. *Journal of Political Economy* **81**(4), 813–846.

Benhabib, J., J. Perla, and C. Tonetti: 2014, ‘Catch-up and Fall-back through Innovation and Imitation’. *Journal of Economic Growth* **19**(1), 1–35.

Bennett, L. M. and H. Gadlin: 2012, ‘Collaboration and team science: from theory to practice’. *Journal of Investigative Medicine* **60**(5), 768–775.

Berkes, E. and R. Gaetani: 2019, ‘The Geography of Unconventional Innovation’. Working Paper 3423143, Rotman School of Management Working Paper.

Bikard, M., F. Murray, and J. S. Gans: 2015, ‘Exploring Trade-offs in the Organization of Scientific Work: Collaboration and Scientific Reward’. *Management Science* **61**(7), 1473–1495.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb: 2020, ‘Are Ideas Getting Harder to Find?’. *American Economic Review* **110**(4), 1104–44.

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

Bloom, N. and J. Van Reenen: 2002, ‘Patents, Real Options and Firm Performance’. *The Economic Journal* **112**(478), C97–C116.

Borjas, G. J.: 1987, ‘Self-Selection and the Earnings of Immigrants’. *American Economic Review* **77**(4), 531–553.

Borjas, G. J. and K. B. Doran: 2012, ‘The Collapse of the Soviet Union and the Productivity of American Mathematicians*’. *Quarterly Journal of Economics* **127**(3), 1143–1203.

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio, and S. J. Terry: 2020, ‘Immigration, Innovation, and Growth’. Working Paper 27075, National Bureau of Economic Research.

Choo, E. and A. Siow: 2006, ‘Who Marries Whom and Why’. *Journal of Political Economy* **114**(1), 175–201.

Ciccone, A. and E. Papaioannou: 2009, ‘Human Capital, the Structure of Production, and Growth’. *The Review of Economics and Statistics* **91**(1), 66–82.

Costinot, A.: 2009, ‘On the origins of comparative advantage’. *Journal of International Economics* **77**(2), 255 – 264.

Davis, S. J.: 1997, ‘Sorting, Learning, and Mobility When Jobs Have Scarcity Value: A Comment’. *Carnegie-Rochester Conference Series on Public Policy* **46**, 327–338.

De Dreu, C. and L. Weingart: 2003, ‘Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis’. *Journal of Applied Psychology* **88**(4), 741–749.

De Dreu, C. K.: 2006, ‘When Too Little or Too Much Hurts: Evidence for a Curvilinear Relationship Between Task Conflict and Innovation in Teams’. *Journal of Management* **32**(1), 83–107.

de Wit, F., L. Greer, and K. Jehn: 2011, ‘The Paradox of Intragroup Conflict: A Meta-Analysis’. *Journal of Applied Psychology* **97**, 360–90.

Eeckhout, J. and B. Jovanovic: 2002, ‘Knowledge Spillovers and Inequality’. *American Economic Review* **92**(5), 1290–1307.

Falk-Krzesinski, H. J., K. Borner, N. Contractor, S. M. Fiore, K. L. Hall, J. Keyton, B. Spring, D. Stokols, W. Trochim, and B. Uzzi: 2010, ‘Advancing the Science of Team Science’. *Clinical and Translational Science* **3**(5), 263–266.

Forman, C. and N. van Zeebroeck: 2012, ‘From Wires to Partners: How the Internet Has Fostered R&D Collaborations Within Firms’. *Management Science* **58**(8), 1549–1568.

Freeman, R. and W. Huang: 2015, ‘Collaborating with People Like Me: Ethnic Coauthorship within the United States’. *Journal of Labor Economics* **33**(S1), S289 – S318.

Freeman, R. B., I. Ganguli, and R. Murciano-Goroff: 2014, *Why and Wherefore of Increased Scientific Collaboration*, pp. 17–48. University of Chicago Press.

Galichon, A., S. D. Kominers, and S. Weber: 2019, ‘Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility’. *Journal of Political Economy* **127**(6), 2875–2925.

Garicano, L. and E. Rossi-Hansberg: 2006, ‘Organization and Inequality in a Knowledge Economy*’. *The Quarterly Journal of Economics* **121**(4), 1383–1435.

Grossman, G. M. and E. Rossi-Hansberg: 2008, 'Trading Tasks: A Simple Theory of Offshoring'. *American Economic Review* **98**(5), 1978–97.

Guimera, R., B. Uzzi, J. Spiro, and L. Amaral: 2005, 'Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance'. *Science (New York, N.Y.)* **308**, 697–702.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg: 2001, 'The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools'. *National Bureau of Economic Research Working Paper no:8498*.

Hayek, F. A.: 1945, 'The Use of Knowledge in Society'. *The American Economic Review* **35**(4), 519–530.

Henderson, R., A. Jaffe, and M. Trajtenberg: 1998, 'Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965Ð1988'. *Review of Economics and Statistics* **80**(1), 119–127.

Hoever, I., D. Knippenberg, W. Ginkel, and H. Barkema: 2012, 'Fostering Team Creativity: Perspective Taking as Key to Unlocking Diversity's Potential'. *The Journal of applied psychology* **97**, 982–96.

Hong, L. and S. E. Page: 2004, 'Groups of diverse problem solvers can outperform groups of high-ability problem solvers'. *Proceedings of the National Academy of Sciences* **101**(46), 16385–16389.

Hunt, J. and M. Gauthier-Loiselle: 2010, 'How Much Does Immigration Boost Innovation?'. *American Economic Journal: Macroeconomics* **2**(2), 31–56.

Jaffe, A. and M. Trajtenberg: 1999, 'International Knowledge Flows: Evidence From Patent Citations'. *Economics of Innovation and New Technology* **8**(1-2), 105–136.

Jaimovich, N. and S. Rebelo: 2017, 'Nonlinear Effects of Taxation on Growth'. *Journal of Political Economy* **125**(1), 265–291.

Jang, S.: 2017, 'Cultural Brokerage and Creative Performance in Multicultural Teams'. *Organization Science* **28**, 993–1009.

Jaravel, X., N. Petkova, and A. Bell: 2018, 'Team-Specific Capital and Innovation'. *American Economic Review* **108**, 1034–1073.

Jehn, K. A.: 1995, 'A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict'. *Administrative Science Quarterly* **40**(2), 256–282.

Jehn, K. A.: 1997, 'A Qualitative Analysis of Conflict Types and Dimensions in Organizational Groups'. *Administrative Science Quarterly* **42**(3), 530–557.

Jones, B. F.: 2009, 'The Burden of Knowledge and the "Death of the Renaissance Man" Is Innovation Getting Harder?'. *The Review of Economic Studies* **76**(1), 283–317.

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. and J. Williams: 1998, 'Measuring the Social Return to R&D'. *The Quarterly Journal of Economics* **113**(4), 1119–1135.

Jones, C. I.: 2005, 'Growth and Ideas'. Vol. 1 of *Handbook of Economic Growth*. Elsevier, pp. 1063 – 1111.

Kerr, W. and W. Lincoln: 2010, 'The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention'. *Journal of Labor Economics* **28**(3), 473–508.

Kerr, W. R.: 2007, 'The Ethnic Composition of US Inventors'. Harvard Business School Working Papers 08-006, Harvard Business School.

Kerr, W. R.: 2018, *The Gift of Global Talent*. Stanford University Press.

Kline, P., N. Petkova, H. Williams, and O. Zidar: 2019, 'Who Profits from Patents? Rent-Sharing at Innovative Firms*'. *The Quarterly Journal of Economics* **134**(3), 1343–1404.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman: 2017, 'Technological Innovation, Resource Allocation, and Growth'. *The Quarterly Journal of Economics* **132**(2), 665–712.

Konig, M., J. Lorenz, and F. Zilibotti: 2016, 'Catch-up and Fall-back through Innovation and Imitation'. *Theoretical Economics* **11**(3), 1053–1102.

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

Lindenlaub, I.: 2017, 'Sorting Multidimensional Types: Theory and Application'. *The Review of Economic Studies* **84**(2), 718–789.

Lucas, R. E.: 1988, 'On the Mechanics of Economic Development'. *Journal of Monetary Economics* **22**(1), 3 – 42.

Lucas, R. E.: 2009, 'Ideas and Growth'. *Economica* **76**(301), 1–19.

Lucas, R. E. and B. Moll: 2014, 'Knowledge Growth and the Allocation of Time'. *Journal of Political Economy* **122**(1), 1—51.

Luttmer, E. G. J.: 2015, 'An Assignment Model of Knowledge Diffusion and Income Inequality'. Staff Report 509, Federal Reserve Bank of Minneapolis.

McFadden, D.: 1974, 'Conditional Logit Analysis of Qualitative Choice Behavior,'. Frontiers in Economics, Chapter 4, ed. d. by P. Zarembka, New York: Academic Press.

Moser, P., A. Voena, and F. Waldinger: 2014, 'German Jewish Emigres and US Invention'. *American Economic Review* **104**(10), 3222–55.

Page, S.: 2007, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press.

Perla, J. and C. Tonetti: 2014, 'Equilibrium Imitation and Growth'. *Journal of Political Economy* **122**(1), 52—76.

Porter, M. E. and S. Stern: 2001, 'Innovation: Location Matters'. *MIT Sloan Management Review* **42**(4), 28–36.

Porzio, T.: 2017, 'Cross-Country Differences in The Optimal Allocation of Talent and Technology'. Working paper.

Romer, P.: 1990, 'Endogenous Technological Change'. *Journal of Political Economy* **98**(5), S71–102.

Singh, J. and L. Fleming: 2010, 'Lone Inventors as Sources of Breakthroughs: Myth or Reality?'. *Management Science* **56**(1), 41–56.

Stokey, N. L.: 2018, 'Technology, Skill, and the Wage Structure'. *Journal of Human Capital* **12**(2), 343–384.

Tao, T.: 2017, 'Australian Spotlight Interview, minute 20:55'.

Teodoridis, F.: 2018, 'Understanding Team Knowledge Production: The Interrelated Roles of Technology and Expertise'. *Management Science* **64**(8), 3625–3648.

Waldinger, F.: 2016, 'Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge'. *Review of Economics and Statistics* **98**(5), 811–831.

Wu, L., D. Wang, and J. A. Evans: 2019, 'Large teams develop and small teams disrupt science and technology'. *Nature* **566**(7744), 378–382.

Wuchty, S., B. F. Jones, and B. Uzzi: 2007, 'The Increasing Dominance of Teams in Production of Knowledge'. *Science* **316**(5827), 1036–1039.

A Theoretical Appendix

This theoretical appendix contains proofs of Propositions 1, 2, and 3.

A.1 Proposition Proofs

Proof of Proposition 1

Proof. I ensure there exists an equilibrium that satisfies the 5 conditions set out in the equilibrium definition in Section 3. For notational purposes, I treat distinct types of skills x and discard location for notational purposes.

The mass of type x that is assigned to team k follows from the mass M_x multiplied by the probability a type x goes to team k. Following McFadden (1974), I show how I derive this probability given the set of sharing rules.

$$Pr\{k_x^*(i) = k\} = Pr(w_k^x + \epsilon_k^x(i) > w_{\tilde{k}}^x + \epsilon_{\tilde{k}}^x(i) \;\; \forall \;\; \tilde{k} \neq k)$$

I take individuals' shocks across all teams k as $F(\epsilon) = \exp(\exp(-\epsilon/\phi))$. Then, with utility from team k as follows:

$$\pi_k^x(i) = w_k^x + \epsilon_k^x(i)$$

The probability individual $i \in x$ chooses team k is:

$$\mathbb{P}\{\epsilon_{\tilde{k}}^x(i) < w_k^x - w_{\tilde{k}}^x + \epsilon_k^x(i) \; \forall \; \tilde{k} \neq k\}$$

$$= \int_{-\infty}^{\infty} \prod_{k \neq k} F(w_k^n - w_{\tilde{k}}^n + \epsilon_k^n) f(\epsilon_k^n) d\epsilon_k^n$$

Plug in the distribution of the shocks

$$= \int_{-\infty}^{\infty} \prod_{\tilde{k} \neq k} \exp\{-\exp[-(w_k^x - w_{\tilde{k}}^x + \epsilon_k^x)]\} \exp[-\epsilon_k^x - \exp(\epsilon_k^x)] d\epsilon_k^x$$

I perform a change of variable to generate $\psi = \exp(-\epsilon_k^x)$ and $z_{\tilde{k}} = \exp[-(w_k^x - w_{\tilde{k}}^x)]$. then:

$$\mathbb{P}(k^* = k) = \int_0^\infty \exp\left[\psi\left(1 + \sum_{\tilde{k}\neq k} z_{\tilde{k}}\right)\right] d\psi = \frac{1}{1 + \sum_{\tilde{k}\neq k} z_{\tilde{k}}} = \frac{\exp(w_k^n/\phi)}{\sum_{\tilde{k}\in\mathcal{T}_n}\exp(w_{\tilde{k}}^n/\phi)}$$

Optimization leads us to the assignment to team k as follows:

$$m_k^x = M_x \cdot \frac{\exp M_{x,\ell}(w_k^x/\phi)}{\sum_{\tilde{k}\in\mathcal{T}_x}\exp(w_{\tilde{k}}^x/\phi)} \tag{10}$$

I use the knowledge of the value of working alone ($q_{\underline{k}}(x)$):

$$m_{\underline{k}} = M_x \frac{\exp(q_{\underline{k}}(x)/\phi)}{\sum_{\tilde{k}\in\mathcal{T}_x}\exp(w_{\tilde{k}}^x/\phi)}$$

as well as the market clearing condition in teams,

$$m_k^x = N_k^x m_k$$

To simplify Equation (10) as follows:

$$\log m_k - \log\frac{m_{\underline{k}}}{N_k^x} = \frac{w_k^x - q_{\underline{k}}(x)}{\phi}$$

Finally, I use Equation (E2) ($\sum_k w_k^{x'} = q_k - c_k$) to sum up this equation across each agent in the team, to get:

$$\log m_k - \frac{1}{T}\sum_{x\in k}\log\frac{m_{\underline{k}}}{N_k^x} = \frac{q_k - \sum_{x\in k} N_k^x q_{\underline{k}}(x) - c_k}{\phi T_k}$$

Satisfying (E1)-(E5) delivers an allocation and set of sharing rules that confirms the proposition.

□

Proof of Proposition 2

Proof. Define $\tilde{q}_k = \frac{q_k - \sum_{x\in k} N_k^x q_{\underline{k}}(x) - c_k}{\phi T_k}$. For each type, there is the market clearing condition:

$$M_x = (1+\exp\tilde{q}_{xx})m_{\underline{k}} + \sum_{y\neq x} \exp\tilde{q}_{xy}\sqrt{m_{\underline{k}}m_{y0}}$$

By the method of *displacement*. Take the derivative with respect to $m_{\underline{k}}$ and then plug in for $\exp\tilde{q}$ and m. This delivers the matrix as follows:

$$\begin{bmatrix} \frac{m_{\underline{k}}+m_{xx}+M_x}{2m_{\underline{k}}} & \dots & \frac{m_{xy}}{2m_{\underline{k}}} & \dots & \frac{m_{xX}}{2m_{\underline{k}}} \\ \vdots & \dots & \dots & \dots & \vdots \\ \frac{m_{xX}}{2m_{\underline{k}}} & \dots & \frac{m_{Xy}}{2m_{\underline{k}}} & \dots & \frac{m_{\underline{k}}+m_{XX}+M_x}{2m_{\underline{k}}} \end{bmatrix} d\boldsymbol{m} = d\boldsymbol{M}$$

The key result to show is that the matrix is invertible. If it is invertible, there is a unique distribution of teams that would be reached for a small change in the distribution of types. This follows from the diagonally dominant matrix theorem and the fact that if a matrix is invertible its transpose is invertible. All that is needed is that $m_{\underline{k}} + m_{xx} + M_x > \sum_{y\neq x} m_{xy}$. This immediately follows from the market clearing condition ($\sum_{y\neq x} m_{xy} + m_{\underline{k}} + m_{xx} = M_x$) and the fact that all possible teams are realized ($m_{\underline{k}} > 0 \Rightarrow \sum_{y\neq x} m_{xy} < M_x$) □

Addendum: Identification condition for multi-person teams

$$M_x = m_{\underline{k}} + \sum_{\mathcal{T}_x} N_k^x m_k \quad ; \quad m_k = \exp\tilde{q}_k \prod_{\tilde{x}\in k}\left(\frac{m_{\tilde{x}0}}{N_k^{\tilde{x}}}\right)^{N_k^{\tilde{x}}/T}$$

Goal is to get the components of $m_{\underline{k}}$ that contribute to M_x, and as such:

$$M_x = m_{\underline{k}} + \sum_{\mathcal{T}_x} N_k^x \exp\tilde{q}_k \prod_{\tilde{x}\in k}\left(\frac{m_{\tilde{x}0}}{N_k^{\tilde{x}}}\right)^{N_k^{\tilde{x}}/T}$$

Take derivative:

$$\frac{\partial M_x}{\partial m_{\underline{k}}} = 1 + \sum_{\mathcal{T}_x} \exp\tilde{q}_k \frac{N_k^x}{T}\left(\frac{m_{\underline{k}}}{N_k^x}\right)^{\frac{N_k^x - T_k}{T}} \prod_{\tilde{x}\neq x\in k}\left(\frac{m_{\tilde{x}0}}{N_k^{\tilde{x}}}\right)^{N_k^{\tilde{x}}/T}$$

Plug back in the endogenous components:

$$\frac{\partial M_x}{\partial m_{\underline{k}}} = 1 + \sum_{\mathcal{T}_x} \frac{m_k}{\prod_{\tilde{x}\in k}\left(\frac{m_{\tilde{x}0}}{N_k^{\tilde{x}}}\right)^{N_k^{\tilde{x}}/T}} \frac{N_k^x}{T}\left(\frac{m_{\underline{k}}}{N_k^x}\right)^{\frac{N_k^x - T_k}{T}} \prod_{\tilde{x}\neq x\in k}\left(\frac{m_{\tilde{x}0}}{N_k^{\tilde{x}}}\right)^{N_k^{\tilde{x}}/T_k}$$

$$\frac{\partial M_x}{\partial m_{\underline{k}}} = 1 + \sum_{\mathcal{T}_x} m_k \frac{N_k^x}{T}\left(\frac{m_{\underline{k}}}{N_k^x}\right)^{\frac{N_k^x - T}{T}} \left(\frac{m_{\underline{k}}}{N_k^x}\right)^{-N_k^x/T}$$

$$\frac{\partial M_x}{\partial m_{\underline{k}}} = 1 + \sum_{\mathcal{T}_x} \frac{(N_k^x)^2}{T}\frac{m_k}{m_{\underline{k}}}$$

$$\frac{\partial m_{\tilde{x}}}{\partial m_{\underline{k}}} = \sum_{\mathcal{T}_{x\cup\tilde{x}}} \frac{N_k^x N_k^{\tilde{x}}}{T_k}\frac{m_k}{m_{\underline{k}}}$$

This delivers a condition under which the matrix of interest would be invertible due to diagonally dominant theorem:

$$1 + \sum_{\mathcal{T}_x} \frac{(N_k^x)^2}{T}\frac{m_k}{m_{\underline{k}}} > \sum_{\tilde{x}\neq x}\sum_{\mathcal{T}_{x\cup\tilde{x}}} \frac{N_k^x N_k^{\tilde{x}}}{T}\frac{m_k}{m_{\underline{k}}}$$

This identification condition is met in the quantitative exercise.

Proof of Proposition 3

Proof. Conditional on an agent $i \in (x,\ell)$ choosing a team k, the expected utility of this agent is:

$$\begin{aligned}
\mathbb{E}\left[\pi_k^{x,\ell} | k = \arg\max_{k'\in\mathcal{T}_{x,\ell}} \pi_{k'}^{x,\ell}\right] &= w_k^{x,\ell} + \mathbb{E}\left[\epsilon_k^{x,\ell}(i) \middle| w_k^{x,\ell} + \epsilon_k^{x,\ell}(i) > w_{\tilde{k}}^{x,\ell} + \epsilon_{\tilde{k}}^{x,\ell}(i)\ \forall\ \tilde{k}\neq k\right] \\
&= w_k^{x,\ell} + \mathbb{P}\left[\pi_k^{x,\ell} \middle| k = \arg\max_{k'\in\mathcal{T}_{x,\ell}} \pi_{k'}^{x,\ell}\right]^{-1} \times \\
&\quad \int_{-\infty}^{\infty} \epsilon_k^{x,\ell}(i) \exp\left[-\epsilon_k^{x,\ell}(i) - e^{-\epsilon_k^{x,\ell}(i)}\left(1 + \sum_{\tilde{k}\neq k}\eta_{\tilde{k}}\right)\right] d\epsilon_k^{x,\ell}(i)
\end{aligned}$$

Where I define $\eta_{\tilde{k}} \equiv \exp\left[-\left(\epsilon_k^{x,\ell}(i) - \epsilon_{\tilde{k}}^{x,\ell}(i)\right)\right]$ The result above is standard and comes from the assumption on the distribution of the shocks across teams. I now use a standard result from math, which shows that $\int_{-\infty}^{\infty} x\exp(x - \eta e^x)dx = -(c + \log\eta)/\eta$ with c as Euler's constant.

This delivers the following equation:

$$\mathbb{E}\left[\pi_k^{x,\ell} | k = \arg\max_{k' \in \mathcal{T}_{x,\ell}} \pi_{k'}^{x,\ell}\right] = c + \log\left(\sum_{\tilde{k} \in \mathcal{T}_{x,\ell}} \exp(w_{\tilde{k}}^{x,\ell}/\phi)\right) \tag{11}$$

Note this equation is independent of the specific team k, and only depends on the distribution of potential teams for the given inventor. I unite this equation with the equation that governs the demand equation for type (x,ℓ) in team (x,ℓ) as a sole inventor:

$$m_{\underline{k}}^{x,\ell} = M_{x,\ell} \frac{\exp(q_{\underline{k}}(x)/\phi)}{\sum_{\tilde{k} \in \mathcal{T}_{x,\ell}} \exp(w_{\tilde{k}}^{x,\ell}/\phi)}$$

$$\mathbb{E}\left[\pi_k^{x,\ell} | k = \arg\max_{k' \in \mathcal{T}_{x,\ell}} \pi_{k'}^{x,\ell}\right] = c + \log\left(\frac{M_{x,\ell}}{m_{\underline{k}}^{x,\ell}} \exp(q_{\underline{k}}(x)/\phi)\right) = c + \log\frac{M_{x,\ell}}{m_{\underline{k}}^{x,\ell}} + q_{\underline{k}}(x)/\phi \tag{12}$$

This delivers our result:

$$\mathbb{E}[V_{x,\ell}] \propto cons + \underbrace{q_{\underline{k}}(x)}_{\text{output alone}} + \underbrace{\phi \log\left(\frac{M_{x,\ell}}{m_{\underline{k}}^{x,\ell}}\right)}_{\text{concentration in teams}} \tag{13}$$

□

B Data Appendix

This section discusses some basic facts in the data related to teams and technologies that were mentioned in the main text.

B.1 Rising Team Size and Rising Returns to Teams

This section discusses general patterns related to teams in the data, starting with the increases in team size across a wide array of categories. I start by discussing the team size distribution. Figure

B.1 illustrates the team size distribution across the sample which has a nice shape and indicates the overall "small teams" model of innovators.

Figure B.1: Team Size Distribution

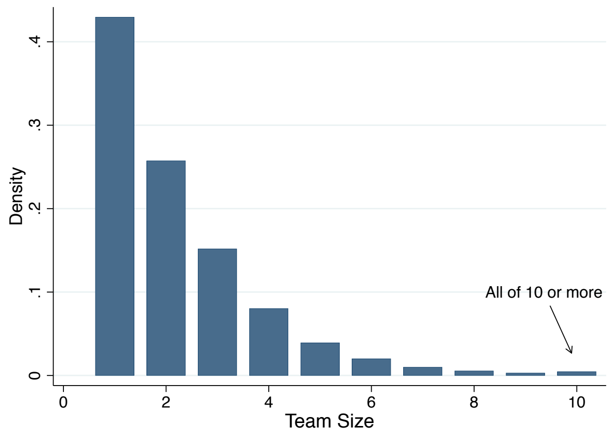
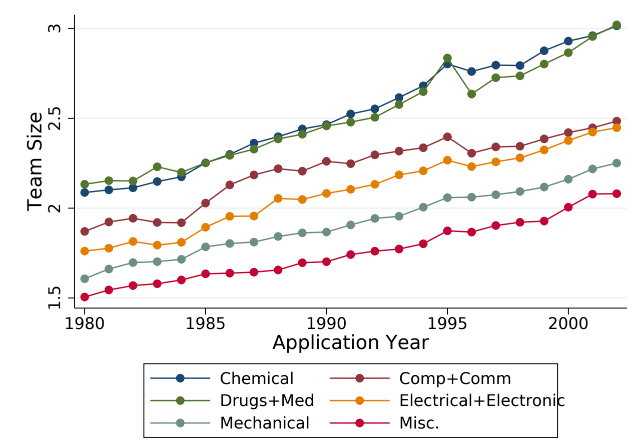


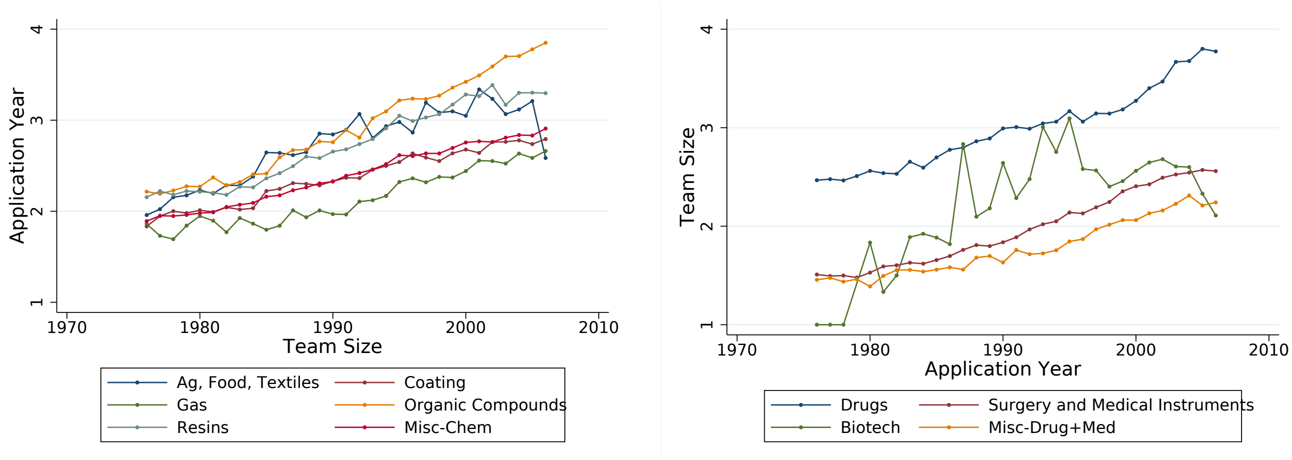
Figure B.2 splits teams by their major category and finds the rise of teams is common across all major technologies. Even splitting by subcategories, as in Figures B.3, B.4, and B.5.

Figure B.2: Mean Team Size by Category



Fact 1 illustrated the rising returns to teams. Here, I direct attention to stock market value instead of citations and find similar results. Figure B.6 illustrates this for stock market value. Figure B.7 adjusts the normalized citations *per person per year* and takes the person's average team size. Thus, this measure accounts for the time it takes to produce a patent. Here, we see that the median optimal team went from 1 in the 1980s to 2 or 3 in the 1990s. Figure B.8 illustrates

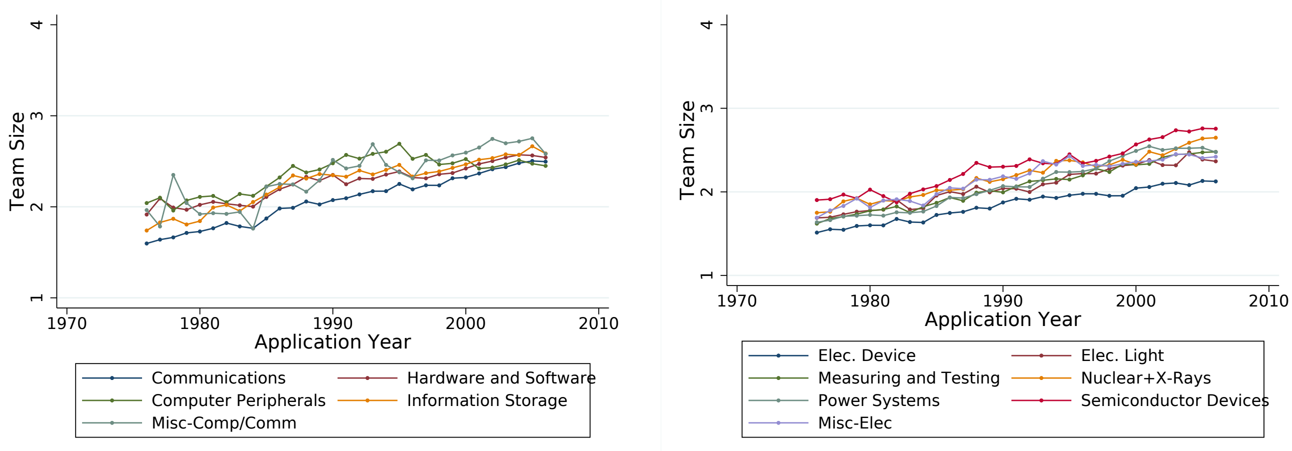
Figure B.3: Team Size Growth by Subcategory



(a) Chemical

(b) Drugs/Medical

Figure B.4: Team Size Growth by Subcategory



(a) Computers/Communications

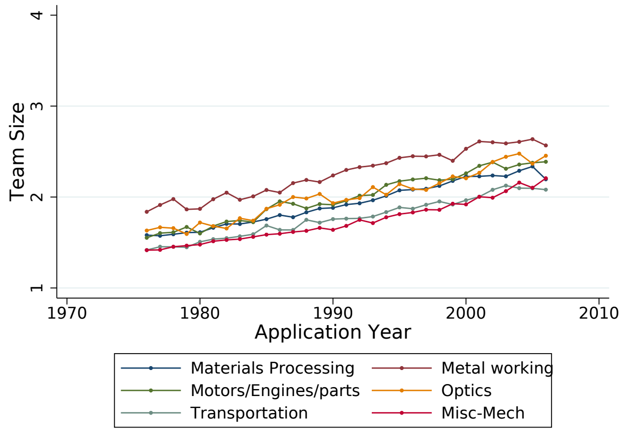
(b) Electrical/Electronics

that the fact comparing the 1980s to the 1990s can be extended to three subperiods, 1976-1985, 1986-1995, and 1996-2005. This plots the mean inverse hyperbolic sine of citations by team size up to teams of size 6. The rising returns to teams appears to be consistent over time.

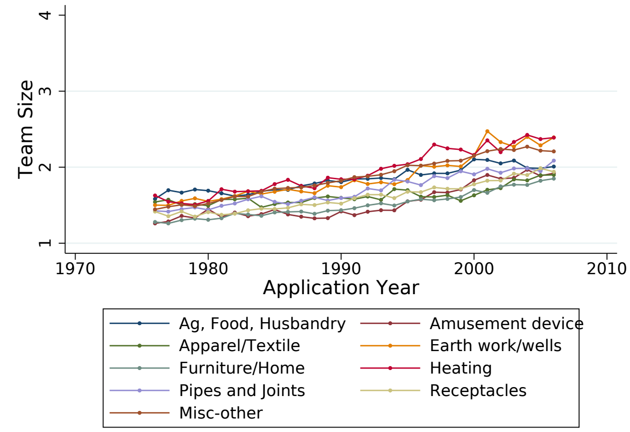
B.2 Rising Diversity

In the introduction, I discussed the rise in team diversity. Here, I document three facts that are indicative of this rising diversity. While the first two (male, female & ethnic diversity) are not relevant for economic output, scholars have noted the link between background diversity and cognitive

Figure B.5: Team Size Growth by Subcategory

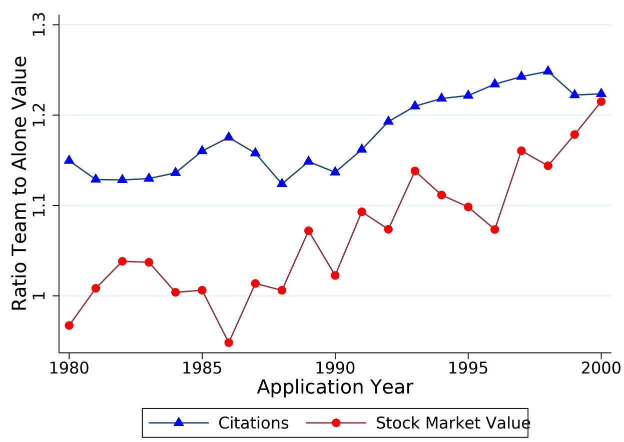


(a) Mechanical



(b) Other

Figure B.6: Citations and Stock Market Value, Team to no Team



diversity (Hoever et al., 2012; Jang, 2017), and found that more diverse teams ethnically produce higher impact patents (see Freeman and Huang, 2015).

I show three graphs that depict the rise in diversity in teams. First, note the rising trend of inventors collaborating together who started in different fields. Figure B.9 illustrates the proportion of patents that have at least two inventors whose initial patent was different from the other inventor. Further, it suggests these are the higher impact patents especially that are being assigned to the diverse teams:

Next, I turn to ethnic diversity. Figure B.10 takes the same ethnic measures as Kerr (2007) and asks how many patents have two distinct ethnicities:

Figure B.7: Citations by Time and Team Size

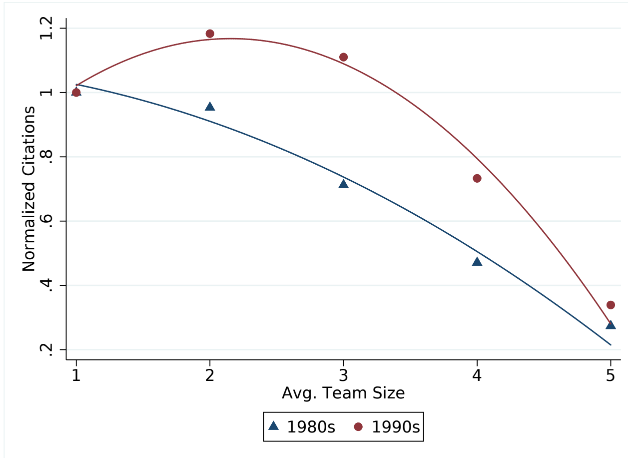


Figure B.8: Split According to Three Periods

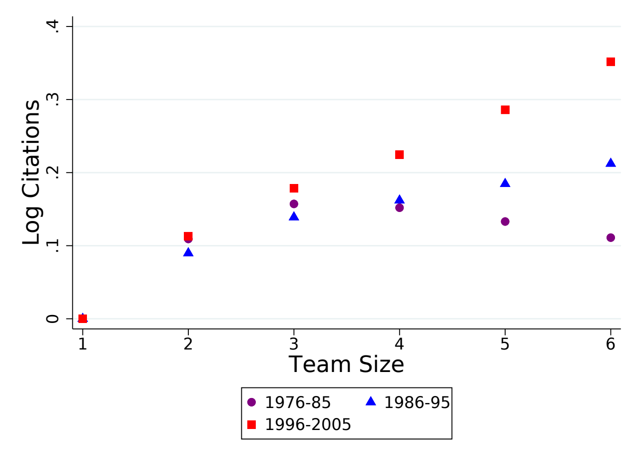


Figure B.9: Prop. patents with ≥ 2 unique technological backgrounds

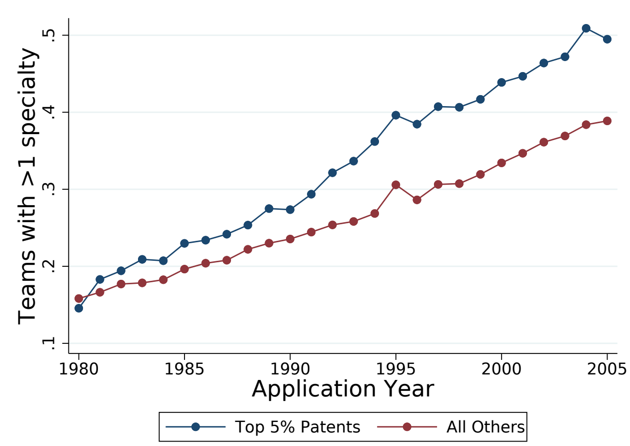
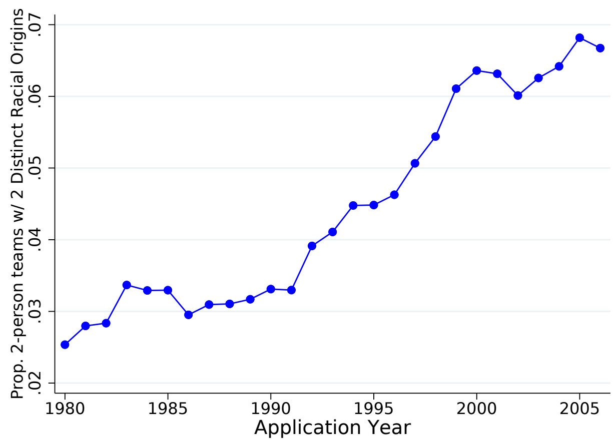
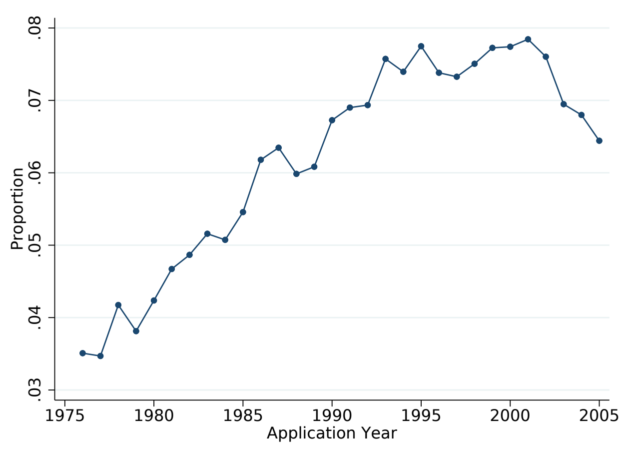


Figure B.10: Prop. 2-person team patents with ≥ 2 unique ethnic backgrounds



Lastly, I note that teams of males and females working together is on the rise, with names probabilistically matched to genders in Figure B.11.

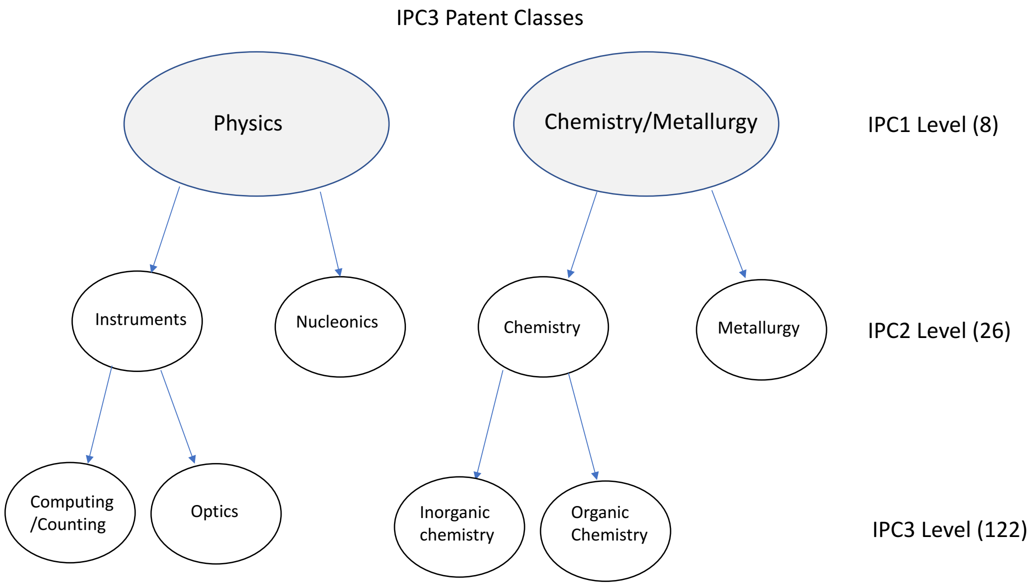
Figure B.11: Proportion of teams with male+female



B.3 Classification Example: IPC3

This paper uses both USPTO and IPC classifications. Both are standard in the patent literature. In order to match the data to records from the Soviet Union, I exploit IPC classification for the quantitative exercise, both at the 2-digit and 3-digit IPC level. Figure B.12 details examples of different layers of classification from IPC1 to IPC3.

Figure B.12: Technology classifications (IPC)



B.4 Unique Inventors and Unique Patents

Table B.1 addresses the count of individuals over time who have had at least one sole-authored patent.

Table B.1: Counts

Time period	Patent Count	Author Count	Avg. Team Size
1980-1990	611103	269286	1.537
1991-2000	968874	361882	1.802
Growth	58.5%	34.3%	**17.2%**

Notes: Patent counts and individuals and team size by decade. Source: USPTO and author calculations

C Empirical Appendix

This section discusses in greater detail the empirical results in terms of both robustness, and expanding on the tests of the quantitative mechanism.

C.1 Expansion on the Empirical Results

Overall, the results in this section on the relationship between patent inputs and outputs provide a framework to link data counterparts to the model to enable quantitative analysis discussed in the next two sections. The specific empirical counterparts discussed in this section are the idea production function which depends on depth and breadth (returns); communication costs from geographical dispersion (costs); and technological and expertise composition (supply).

I start by characterizing the idea production function non-parametrically. Next, I evaluate the role of communication costs in both the quality and frequency of patent production. Lastly, I illustrate that a pure compositional effect of technologies is not a large force in the pattern behind the increase in team size with a simple decomposition. This section produces five empirical results related to the three key forces:

Empirical Result I.a: Patent impact increases significantly with depth and breadth.

Empirical Result I.b: Teams with higher depth and breadth produced even more impactful patents in the 1990s than in the 1980s.

Empirical Result II.a: Geographical distance reduces the probability that a team will form, but its importance diminishes over time.

Empirical Result II.b: Once controlling for inventor expertise, geographical distance does not have strong effects on patent quality.

Empirical Result III: The compositional effects behind changing team size are small.

C.2 Robustness of Empirical Results

When a team produces together within a class, teams apply their depth and breadth to produce patents of quality q_p. In order to adjust for the speed of patent production, I adjust patent quality by the average time it takes to produce the patent across the main specifications. In the following empirical exercises, the goal is to evaluate the relationship between patent value $\log q_p$. For q_p, I take the lifetime-adjusted citations of the focal patent. To ensure that the results are robust, I also use a measure of stock market value from Kogan et al. (2017).

Empirical Result I.a: Patent impact increases significantly with depth and breadth.

I use the following regression to evaluate the response of citations to depth and breadth, and how it has changed.

$$\log q_{p(k,s)} = \alpha_0 + \sum_{j=2}^{10} \alpha_j \mathbb{I}\{Decile(D_{k,s}) = j\} + \sum_{i=2}^{10} \zeta_j \mathbb{I}\{Decile(B_{k,s}) = j\} + \mathbf{Z}_{p(k,s)} + u \quad (14)$$

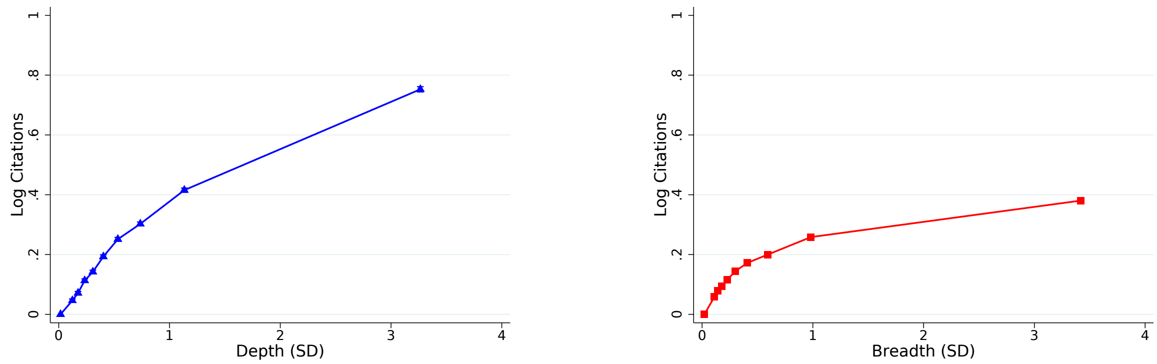
The left-hand side variable in Equation (14) is the log quality ($\log q_p$) of patent p in patent class s produced by team k. The coefficients on the deciles of the depth of team k in technology class s ($D_{k,s}$) and breadth ($B_{k,s}$) are the coefficients of interest, α_j, ζ_j. Due to the richness of the data and collapsing the right hand side into two variables of interest, these nonparametric regressions can be used throughout. $\mathbf{Z}_{p(k,s)}$ contains a set of relevant controls: *patent class* $\times$ *application year*, *team size* (nonparametric), and *team experience* (nonparametric).

β_j is the marginal increase in log citations on patent p in depth decile j versus depth decile 1, holding fixed team size, experience, class$\times$year, and breadth. ζ_j is the marginal increase in log citations on patent p in breadth decile j versus breadth decile 1, holding fixed team size, experience, class$\times$year, and depth. I plot the coefficients and their robust standard errors but detail the results further in the Appendix. Figure C.13 shows the strong association between patent quality and patent depth and breadth. In Figure C.13a, holding team size, team experience, team breadth, and class$\times$year fixed, the 9th decile of depth has approximately 10% more citations than the 8th decile of depth. Similarly, in Figure C.13b, teams in the 9th decile of breadth have around 8% more citations than the 8th breadth decile.

Empirical Result I.b: Teams with higher depth and breadth produced even more impactful patents in the 1990s than in the 1980s.

To evaluate the regression with depth and breadth as the driving force behind production, I

Figure C.13: Log Lifetime Citations and Depth/Breadth



(a) Depth

(b) Breadth

perform the following regression:

$$\log q_{p(k,s)} = \alpha_0 + \sum_{j=2}^{10} \alpha_j \mathbb{I}\{Dec.(D_{k,s}) = j\} \times 1990s + \sum_{i=2}^{10} \zeta_j \mathbb{I}\{Dec.(B_{k,s}) = j\} \times 1990s + \mathbf{Z}_{p(k,s)} + u \tag{15}$$

Equation (15) looks like Equation (14), but it includes an interaction term for the 1990s. The goal is to outline how log citations respond to depth and breadth across the two different time periods. If there were a change to the idea production function over time, then one would expect higher coefficients on depth and breadth in the 1990s relative to the 1980s.

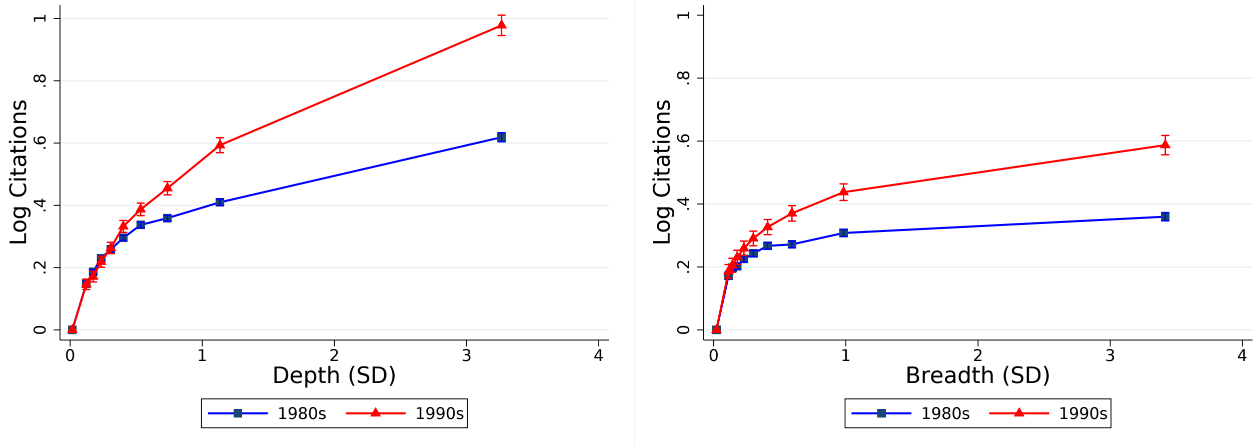
Indeed, as indicated in Figure C.14 by the relationship between the red and blue lines, higher-decile depth and breadth are linked to higher-quality patents in the later period. These results provide evidence that the relative returns to teams are increasing and that the increasing returns come from team skills, as expressed in depth and breadth.

This section characterizes the gross output measures for various collections of skills of teams, q_k. In the next section, I explore the costs of forming teams, c_k, through the lens of inventor geography.

C.3 Communication Costs

This section explores how communication costs affect team formation and patent quality. Given the significant evidence that face-to-face interactions reduce frictions, I use geographical dispersion as a measure of communication costs. I focus on two measures of distance: i) the log of the total

Figure C.14: Citations and Depth/Breadth, 1980s versus 1990s



(a) Depth

(b) Breadth

distance between individuals on a team, and ii) the number of unique locations on the patent. This section discusses the effect of being in different locations on team production and team formation, and further explores the role of distance in Section 5.

Empirical Result II.a: Teams are becoming more geographically dispersed

Geographical distance affects team formation, although over time this effect has become weaker. This was seen in Figure 3.

Empirical Result II.b: Geographical distance and patent quality

The following regression determines whether geographical distance affects patent quality. Once the full set of controls are included, locational distance of inventors on a patent does not affect patent quality. In the following regression, The value of an idea $\log q_p$ is regressed on whether a team is in the same location or a different location, restricting attention to two-person teams so as not to pick up team size effects.

$$\log q_{k(p)} = \alpha_0 + \alpha_1 Diff\_Loc_{k(p)} + \mathbf{Z}_{k(p)} + u \tag{16}$$

Table C.3 finds a *positive* effect of being in a different location on patent quality in column (1). This is sensible as teams that form more often despite high costs will have higher q_k. Once we control for the technology class, depth, and breadth, it appears locational distance does not have an effect on patent quality conditional on the same skill input.

Table C.2: Patent Quality and Different Location

	(1)	(2)
	Log Patent Quality	Log Patent Quality
Different Location	0.052**	0.017
	(0.019)	(0.009)
Class, Depth, and Breadth Control		X
Observations	525043	525043
R^2	0.005	0.212

Note: Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

Compositional Effects on Team Size

I present a simple decomposition exercise to evaluate the contribution of mechanical changes in the distribution of technologies to the increase in team size. Section 5 discusses the inventor skill composition more directly.

Empirical Result III: Compositional effects on changing team size are small.

The change in team size can be written out as follows, where ω_s is the weight of class s among all patent classes at time t and $T_{s,t}$ is the team size of patent class s at time t.

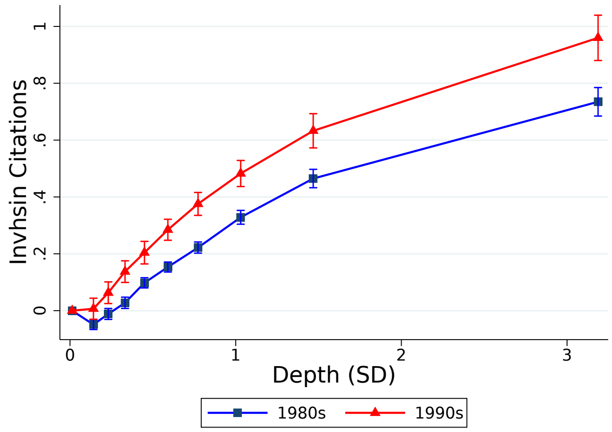
$$\sum_{s\in\mathcal{S}} \underbrace{\omega_{s,2000}T_{s,2000} - \omega_{s,1980}T_{s,1980}}_{\text{change in team size}} = \sum_{s\in\mathcal{S}} \underbrace{(\omega_{s,2000} - \omega_{s,1980})T_{s,1980}}_{\text{compositional change}} + \underbrace{\omega_{s,2000}(T_{s,2000} - T_{s,1980})}_{\text{change within class}}$$

I find that the compositional change explains 7% of the rise in team size, while the change within classes explains the rest.

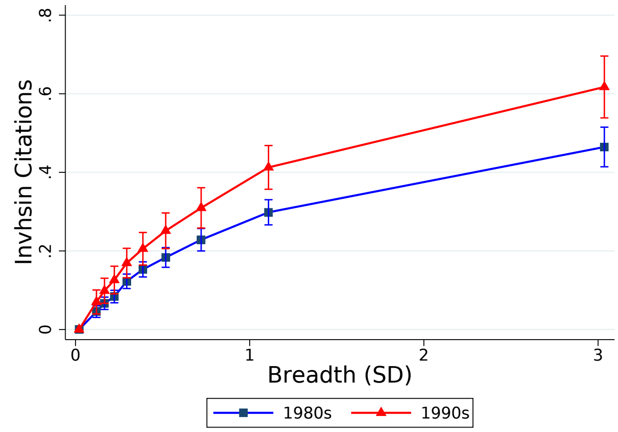
C.4 Idea Production Function

I now turn to a discussion of the mechanics and robustness of the idea production function which takes as inputs depth and breadth. Figure C.15 shows that even when controlling for firm fixed effects, the 1990s shows higher slope response to depth and breadth.

Figure C.15: Citations and Depth/Breadth, Firm Fixed Effects, 1980s versus 1990s



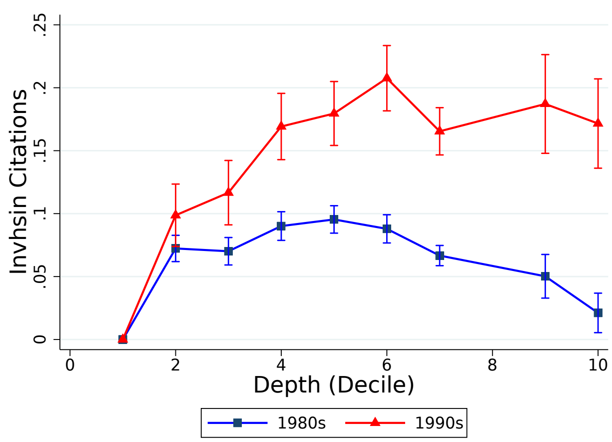
(a) Depth



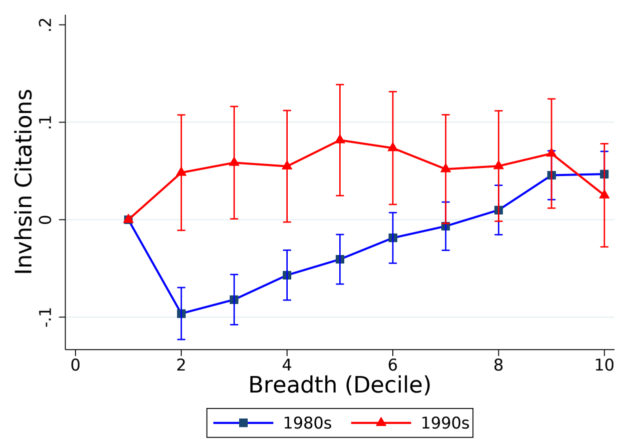
(b) Breadth

Figure C.16 shows that even when controlling for individual fixed effects, the returns to depth and breadth are higher in the 1990s than the 1980s. This is true even when exploring different definitions of breadth.

Figure C.16: Citations and Depth/Breadth, Individual Fixed Effects, 1980s versus 1990s



(a) Depth

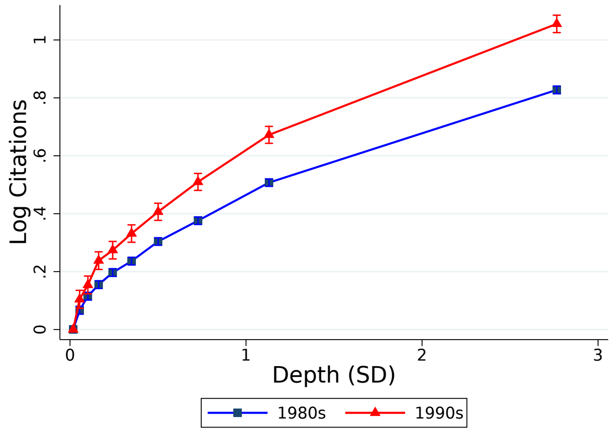


(b) Breadth

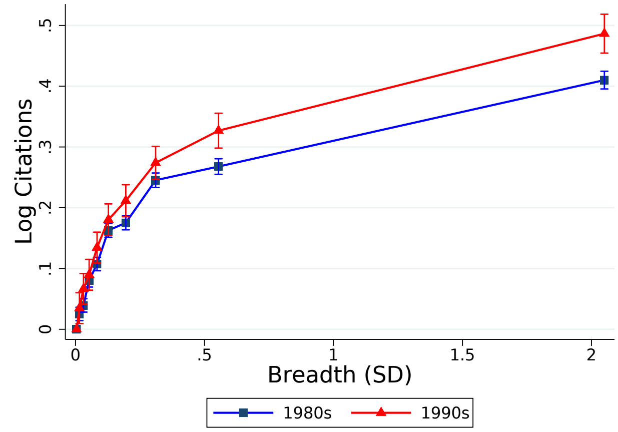
Figure C.17 illustrates that the same result holds when I compute skill only referencing the classes an individual works alone in.

Figure C.18 illustrates that, if anything, the stock market value responsiveness measured by Kogan et al. (2017) has an even stronger response to depth and breadth in the 1990s than the 1980s.

Figure C.17: Use Only Working Alone for Skill, 1980s and 1990s

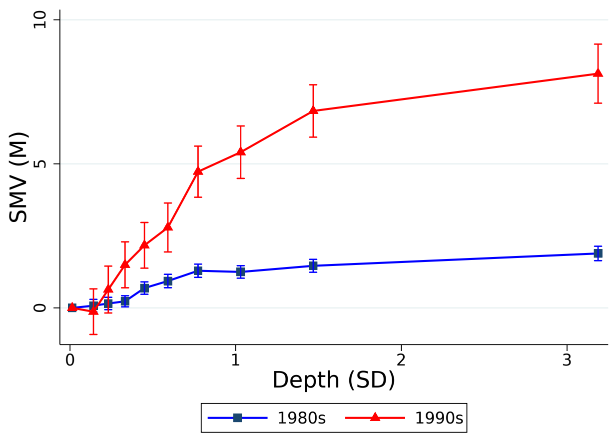


(a) Depth

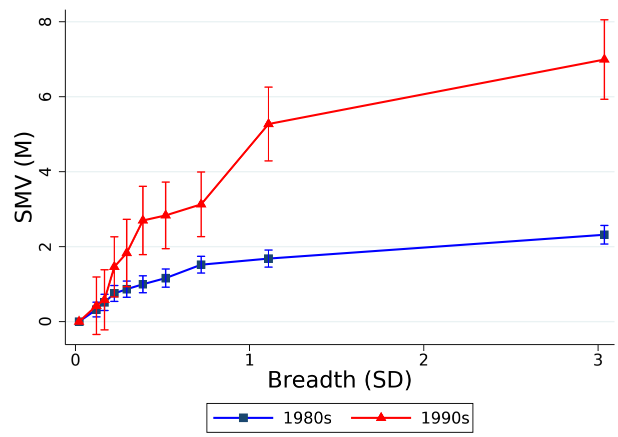


(b) Breadth

Figure C.18: Stock Market Val. and Depth/Breadth, 1980s versus 1990s



(a) Depth

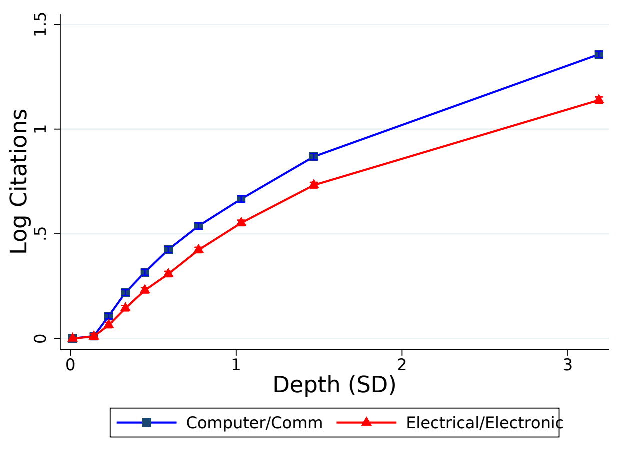


(b) Breadth

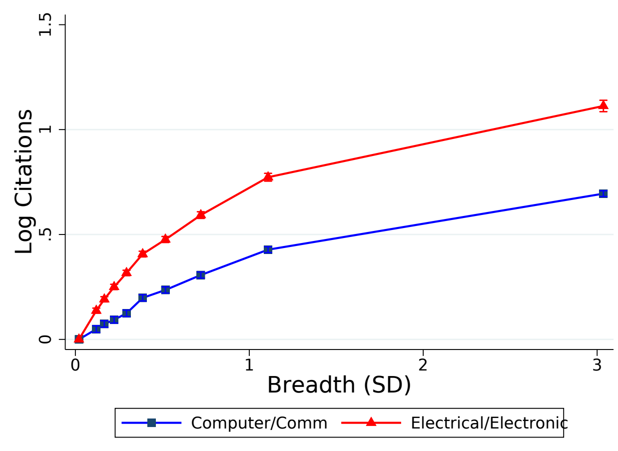
One might expect the returns to depth and breadth to vary due to the underlying technology. I do a split by technologies in order to determine which types of technologies get larger returns to depth and breadth. Figure C.19 demonstrates that electrical processes, which are often more linked to basic research and a broader array of classes, exhibit higher returns to breadth versus Computing and communications. Conversely, computing and communications, which tap into expertise within the focal class, exhibits a larger response to depth than to breadth.

In addition to the heterogeneous responsiveness to depth and breadth, high depth patents are more likely to provide knowledge to the focal technology class through forward citations. High

Figure C.19: Variation across Technologies



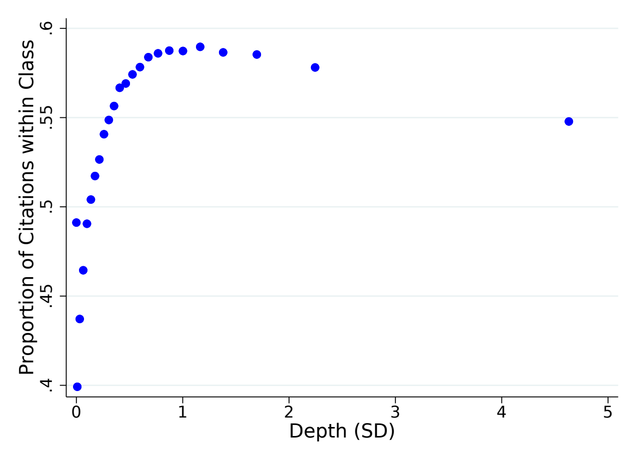
(a) Depth, Electric v. Comp



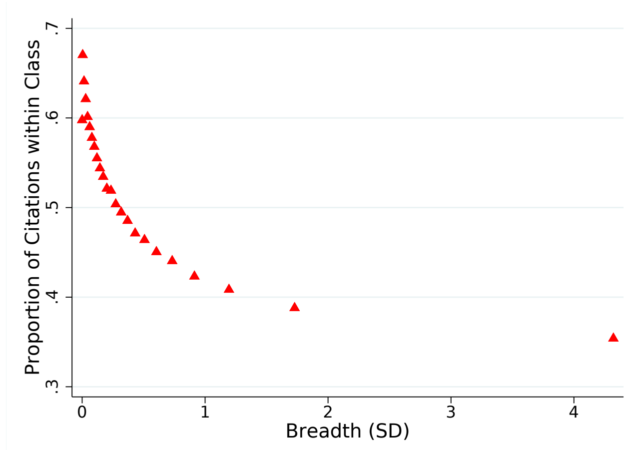
(b) Breadth, Electric v. Comp

breadth patents are more likely to provide knowledge to other technology classes via forward citations. This is true when controlling for a host of other factors like the focal patent class, year, and team size. Figure C.20 illustrates this result:

Figure C.20: Destination Citations



(a) Depth and Forward Citations

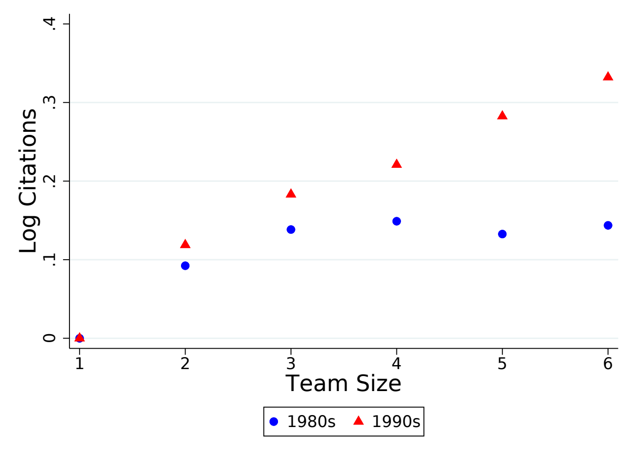


(b) Breadth and Forward Citations

Figure C.21a shows that teams produce patents with higher citations in general, and that this is more pronounced in the 1990s than the 1980s. Figure C.21b shows the increasing returns to teams are better understood through the skills that individuals bring rather than simply team size.

Once I control for depth and breadth on the patent and plot the residuals from these regressions, the relationship between patent quality and team size goes away and does not shift from the 1980s to the 1990s. The fact that controlling for depth and breadth kills the returns to teams indicates the importance of understanding *expertise* for evaluating the changing patent value.

Figure C.21: Citation value and team size, 1980s versus 1990s



(a) Raw Means

(b) +controls for depth and breadth

C.5 Communication Costs

Figure 3 illustrates how more inventors are collaborating across the country. I provide further evidence on the falling communication costs driving this phenomena. Figure C.22 shows the cost of a 10-minute phone call in 2010 US dollars over time. The cost flattens out around $4.00 when email and the internet become more prevalent:

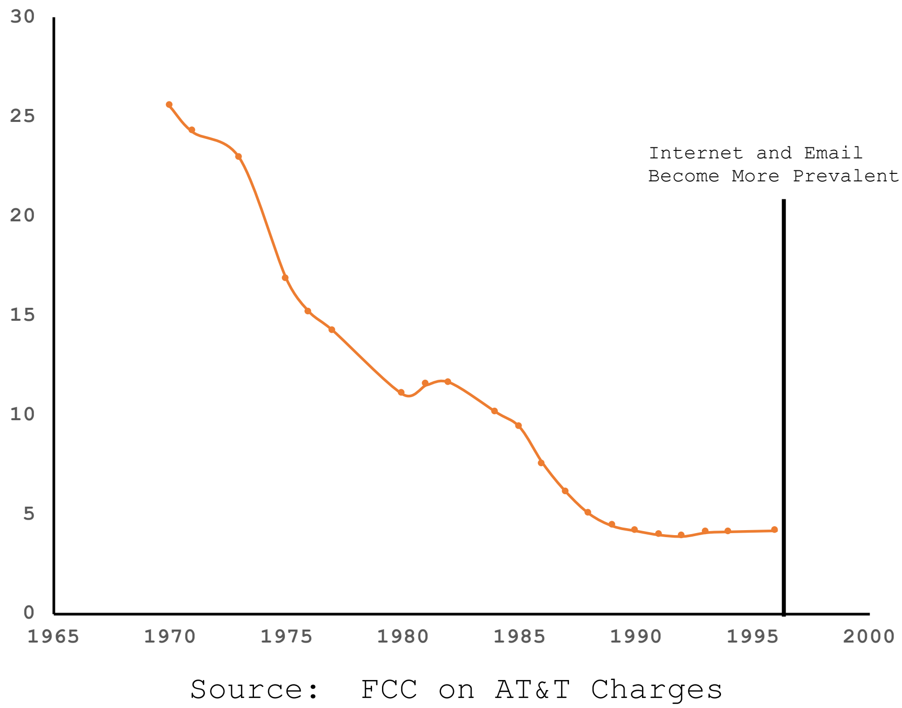
Further, while Figure 3 showed the frequency of two co-authors being in different locations. Here, I show how the relationship is even more stark when one takes distance within teams smaller than four. Figure C.23 plots the mean miles distance by shortest path within the team from 1976-2006.

This section explores how communication costs affect team formation and patent quality. Given the significant evidence that face-to-face interactions reduce frictions, I use geographical dispersion as a measure of communication frictions. Here, I discuss the effect of distance on team formation.

Empirical Result II.a: Teams are becoming more geographically dispersed

Geographical distance affects team formation, although over time this effect has become weaker. The following equation regresses the probability a two-person team is in the same location against

Figure C.22: Cost of a 10-minute cross-country phone call, by year



time.There is a dummy for each decade from the 1970s to the 2000s (with 1970s omitted):

$$Diff\_Loc_p = \alpha_0 + \sum_{j=1980s}^{2000s} \beta_j \mathbb{I}\{Decade = j\} + u \tag{17}$$

While Table C.3 indicates, via the constant term, that only a small proportion of collaborators resided in different locations (12.5%) during the 1970s, in the 2000's there were almost $1.5\times$ as many two-person teams in a different location. If the analysis includes multi-person teams, the results are even more stark.[20] This suggests forming teams at a distance has become easier.

Note from Column (2) in Table C.3 that including a host of controls does not affect the overall pattern. This suggests that the increase in distance between co-authors is not a function of changes in the regional skill distribution or rising skill complementarities. If these increases were a result of rising skill complementarities, then the introduction of the controls would alter the changing pattern of allocations to cross-region teams over time, because it would be the strength of the complementarity driving some of the matching pattern. However, this is not observed. Next, I turn to technology composition effects.

[20] Please see Appendix C.5.

Figure C.23: Geographical Distance on Co-Authored Patents Over Time

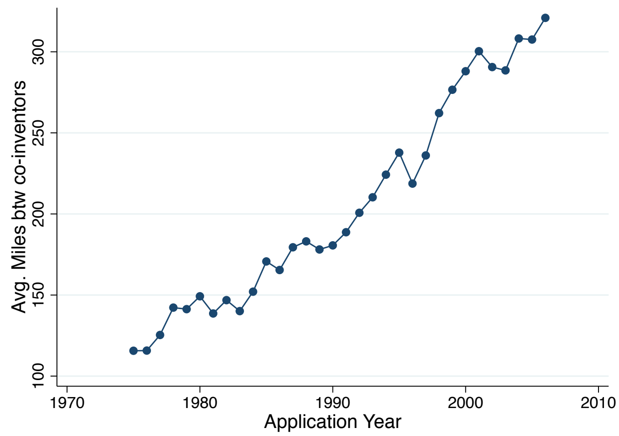


Table C.3: Different Location by Decade

	(1) Different Loc.	(2) Different Loc.
1970s	0	0
	()	()
1980s	0.014	0.013
	(0.002)	(0.002)
1990s	0.036	0.037
	(0.002)	(0.002)
2000s	0.054	0.057
	(0.002)	(0.002)
Constant	0.126	-
	(0.002)	()
Observations	516182	516182
R^2	0.002	0.011
Class/Team Skill Controls	N	Y

Notes: Clustered standard errors at class-level in parentheses

Empirical Result II.b: Geographical distance and patent quality

In addition to showing the distance effect declining over time, I want to ensure the distance itself does not affect patent quality, which enables my primary focus on the link between patent quality and team expertise. Here, I show that locational distribution of inventors on a patent does not affect patent quality. In the following regression, I regress the value of an idea $\log \lambda_p$ on whether a team is in the same location or a different location. The goal of this regression is to isolate the effect of inventor distance on patent quality.

$$\log \lambda_p = \alpha_0 + \alpha_1 Diff\_Loc_p + \mathbf{Z}_{p(\tau,s)} + u \tag{18}$$

I find an *positive* effect of locational distance on patent quality, which is sensible through the lens of the model: teams that form more often despite high costs will have higher λ_τ. Once we control for the technology class, as well as depth and breadth, these forces go away. It appears locational distance does not have an effect on patent quality conditional on the same skill input.

Table C.4: Patent Quality and Different Location

	(1)	(2)
	Log Patent Quality	Log Patent Quality
Different Location	0.052**	0.017
	(0.019)	(0.009)
Class Controls		X
Depth Control		X
Breadth Control		X
Observations	525043	525043
R^2	0.005	0.212

Notes: Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

I also build on Table C.3 to show a similar result when we include distance and teams greater than size two and include individual fixed effects. Table C.5 illustrates this. One can observe in column (1) again the positive effect of log distance on patent quality. This declines with the introduction of controls for breadth and depth in column (2). Once individual fixed effects are controlled for in column (3), this effect goes away.

Table C.5: Patent Quality and Log Distance

	(1) Log Patent Quality	(2) Log Patent Quality	(3) Log Patent Quality
Log Distance	0.00579**	0.00359**	-0.00037
	(0.00017)	(0.00016)	(0.00024)
Breadth and Depth Controls		X	X
Individual Controls			X
Observations	2725471	2725471	2123280
R^2	0.1239	0.1791	0.5632

Notes: Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$

D Policy and Immigration Application

The main text discusses the inflow of Russians from the former Soviet Union and their role in teams. This section addresses more details of the skill distribution of newly arrived Russians and the robustness of the general results. Appendix D.1 discusses the application from the fall of the Soviet Union and specific technologies and the robustness of the results to changes in the categorization of team types and output measures. Appendix D.2 discusses the more general immigration policy results, the role of self-selection in migration amplifying innovation, and the general results on the economic value of specific types. Appendix D.3 discusses R&D and education policy applications. Appendix D.4 discusses the computational exercise of shocking the system. Appendix D.5 discusses the value of specific types in private and general equilibrium.

D.1 Soviet Union Shock

In the main text, I use IPC2 classifications for talent supply shocks. In exploring the robustness, I use a more granular definition of skills by mapping individual expertise to IPC3 classes with sufficient observations (122 unique categories). I take the individual's top category when working alone, then on teams of 2, etc. Inventors with only one patent observation are not classified according to a type.

Figure D.24 plots the concentration of the Soviet Union and the United States across patent classes according to the IPC3 patent classifications. This figure illustrates the heterogeneous exposure across classes that enable use of the fall as a shock to the supply of talent.

Figure D.25 illustrates that the concentration of Russians across types in the IPC3 categories has some resemblance of matching the pre period. However, there is selection in the migration pattern, as discussed in Section 6.

Self-selection amplifies the Russian overall contribution to output. This can be observed in the main text when I use IPC2 types and teams up to size three. Here, we use teams up to size two and types by IPC3 and find similar results.

Table 4 affirms the two main results from the inflow of Russians. First, the Russian contribution to aggregate innovation was more than their increase in the population in terms of inventors (0.93 vs. 0.8). Second, this was due to the self-selection of migrants.

Figure D.24: Concentration of US and Soviet Union across IPC3, 1980-1990

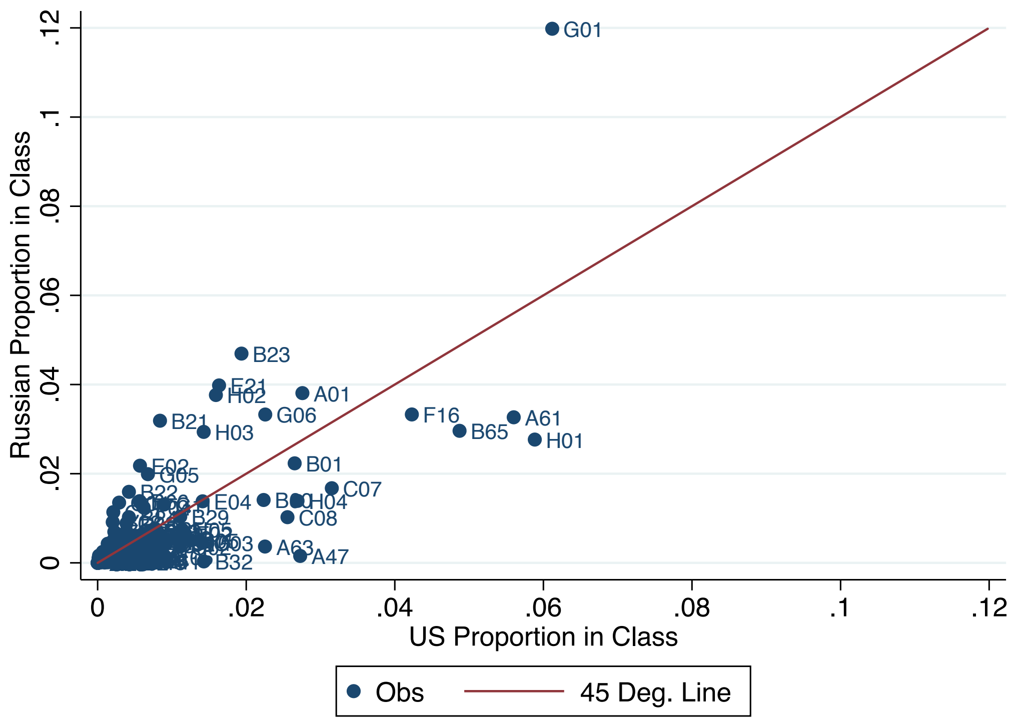


Figure D.25: Concentration of US and Soviet Union across IPC3, 1991-2000

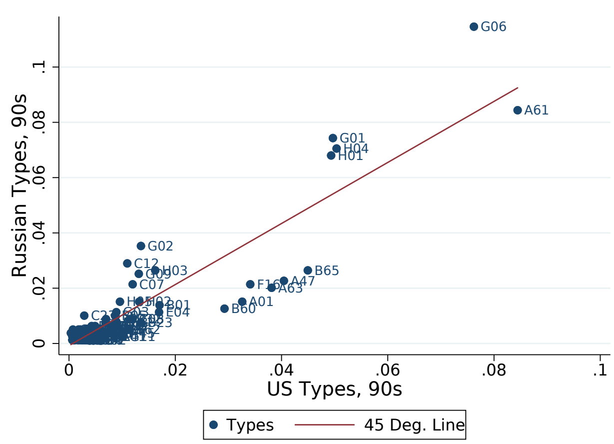


Table D.7 follows from Table D.6 which uses citations, but restricts attention to patents that have an associated stock market value (Kogan et al., 2017). Note again that there is a fairly close match to the predicted and realized in-sample exercise, where the selection on the types arriving to the US shapes the overall contribution. This confirms in yet another measure the value of self-selected migration and its amplification.

Table D.6: Contribution to Aggregate Innovation, 1995-2005

Measure	**Δ Agg. Innov (%)**	***Corr(Pred. SU, RU in US)***
	— Panel A. Predicted Impact, SU shock —	
No selection	0.56	0.51
Selection at T50	0.79	0.62
Selection at T20	1.07	0.66
	— Panel B. Impact of SU shock —	
Russian inflow	0.93	1

Notes: Calibrated model output using counterfactual skill distribution. In Panel A, each row represents a simulated distribution of teams adding the shock from the Soviet talent, with different selection thresholds (e.g. proportional to selection) from Equation (8). In Panel B, the actual Russian distribution is simulated.

Table D.7: Stock market value of migrating Russian output, 1995-2005

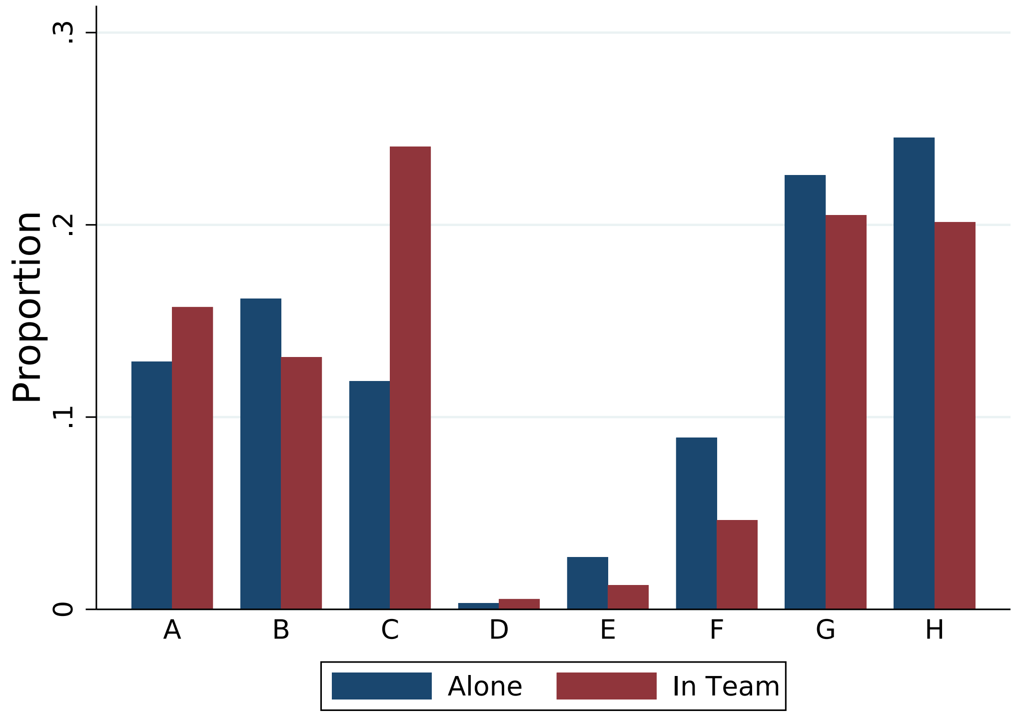
Measure	**Δ Agg. Innov ($)**
— Panel A. Innovation in US —	
sole-authored innovation	10.8B
model predicted (incl. teams)	33.4B
— Panel B. Predicted innovation from SU-US Match —	
predicted, from SU data, no selection	26.2B
predicted, from SU data, selection at T50	32.5B

Notes: Market value of patents (in sample) from 1995-2005 is 3.92T

Figure D.26 illustrates that Russians contributed to certain classes in teams, while they contributed to other classes more often alone. In particular, Russians were not heavily concentrated in chemistry (C) alone, while they were concentrated in classes related to physics and electricity (G,H). Given the fact that chemistry is often pulling in large teams with varied expertise and the Soviet Union had low representation in this field, this is not surprising:

The idea production function and communication costs push in opposite directions to impact change in aggregate innovation that results from immigrant inflows. First, an addition of a worker to a country is of higher value because of their ability to contribute to teams. This would suggest immigration policy is becoming more important and it is crucial to link immigrants into the global

Figure D.26: Russian Innovations Alone and in Teams



market. Second, because international collaboration is increasing, it is less important to bring immigrants directly into the home country. Thus, it is a quantitative question of what is the dominant force to consider when designing immigration policy. We find in the quantitative section the idea production function is the most important component driving the increase in teams. Figure D.27 finds that cross international collaboration is increasing but still not a huge proportion of patents.

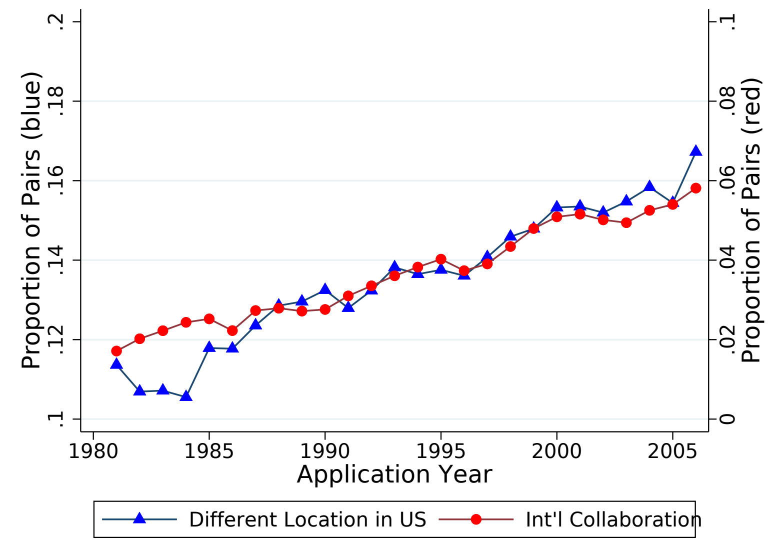
D.2 General Immigration Counterfactuals

To understand both the value of specific types and the correlation between self-selection and aggregate innovation, I perform a quantitative exercise aimed at understanding the impact of each type on innovation. I increase the supply of a specific inventor type in the economy by a small amount and evaluate how this changes overall innovative output.

In these counterfactuals, I take the estimated production function from the 1990s for each q_k, and the supply of each type x, M_x.[21] I use IPC2 classifications to characterize the expertise of

[21] For the case of immigration inflows, I abstract away from regional communication costs within the US.

Figure D.27: Co-Inventors across locations (2-person teams)



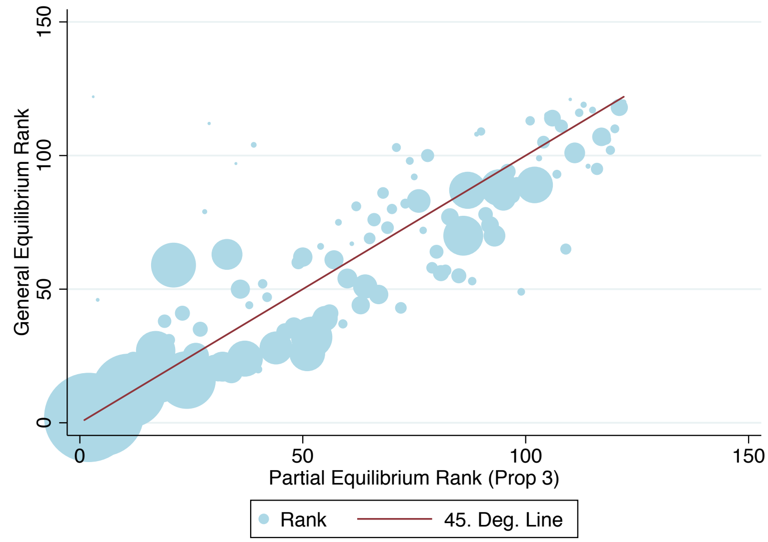
each individual. I use these classifications to match the distribution of types, M_x, and teams, m_k, and team production, q_k, using US data from 1991-2000. The exercise then increases each type M_x by 0.01 with the total population normalized to 1.

Returning to Equation (8), I compare the rank of total output generated by 1% increase in the inventor population of a specific expertise across 26 IPC3 categories to the rank from Proposition 3. Figure D.28 plots the ranks from ex ante value and ex post contribution for the 1990s time period, with the bubble size indicating the mass of each type. Notice the strong relationship between the value from Proposition 3 and the general equilibrium result of shocking the economy:

For each type, two forces – the value produced alone $q_{\underline{k}}(x)$ and the frequency that this type joins a team – are close to sufficient to characterize the outcome of increasing this specific expertise in the economy through immigration. The team element plays a key role in this value, and indicates why we see more migration from experts in Biochemistry, Organic Chemistry, and Medical Science.

I stress two takeaways from this result. First, the self-selection of immigrants has a tendency to amplify the effect of migration on innovation, as immigrants with a better fit of expertise for the society will be more likely to migrate, as was seen in the case of the Soviet Union. Second, for a policymaker whose goal is increasing aggregate innovation, there is only a limited set of information required to understand which types of expertise will make the largest contribution: the

Figure D.28: Relationship between PE values and GE result



Notes: This figure compares ranks of IPC3 types in partial equilibrium (own marginal value) and general equilibrium (own effect on aggregate). Size of dots indicate share of individuals with given expertise.

productivity of types when they work alone ($q_{\underline{k}}(x)$), their concentration in teams, and the noise parameter, ϕ. Even though the general equilibrium forces don't make this a perfect predictor of overall contribution, its ability to approximate the outcome can be informative for immigration policies based on skill.

Policymakers without knowledge of the noise parameter ϕ can rely on two straightforward pieces of information: how productive is a specific inventor type (e.g. organic chemist) alone, and how often are they in teams? Table D.8 illustrates the top expertise in the 1990s with the estimated production function.

D.3 Discussion: R&D Subsidies and Education Policy

The team production channel induces a lot of potentially heterogeneous impact of R&D and education policy depending on its structure. The effect of R&D subsidies in this framework depend on how they target the idea market. For instance, if subsidies are directed to labs (teams), this will naturally change team composition through inducing more collaboration. Agents might be more willing to overcome communication costs in order to join teams. If a team's communication costs are borne privately while team innovative output is a public good, this could increase overall

Table D.8: Ranking Types across IPC3

	(1)	(2)	(3)
Rank (IPC3)	1990s $E[V_x]$ rank	1990s, alone	1990s, in teams
1	Checking-Devices	Checking-Devices	Biochemistry
2	Medical Science	Medical Science	Organic Chemistry
3	Computing/Counting	Computing/Counting	Organic Macromolecular Compounds
4	Elec. Comm. Technique	Elec. Comm. Technique	Fatty Acids
5	Biochemistry	Signaling	Petroleum and Technical Gases

Notes: Comparing ranks of top 5 IPC3 skills and indicating the relative value alone and in teams: Source: USPTO and author calculations.

innovation. If R&D subsidies are directed towards all innovation, it will exacerbate the differences across productive and unproductive teams.

R&D subsidies would not directly hit communication costs. A subsidy on communication technology would induce more regional dispersion of teams. This is part of a natural progression, but a subsidy would induce more of this behavior. However, the R&D subsidy will hit differentially across the quality of teams q_k, whereas communication technology subsidies exclusively hit the high communication cost teams.

A policy of R&D subsidies seems more sensible when immigration policy becomes less feasible. Immigration policy does not require taxation that pulls resources away from other projects. Given that most patents that come from inventors across international borders are organized within firms, the presence of multinational firms would seem to be a useful mechanism for generating this subsidy to communication technology.

I also address the role of education policy in a qualitative manner. Through the lens of the model, education policy follows similar principles to immigration policy. If the cost of training is equal across domains of expertise, then education should be tilted towards classes that have the largest aggregate effects through both their own production channel and the team production channel. Further, fostering diverse expertise will be important for the economy given the rise of diverse expertise in teams.

However, it may be true that the high-value expertises are more costly to train. For instance, organic chemists are extremely valuable because they make contributions across classes. However, it is costly to train organic chemists. This paper has not built a framework to evaluate the mechanisms governing this tradeoff. Because the social value and private value of innovation can be misaligned, this seems like an exciting path for future research.[22]

A key result for education policy is that specific curricula should interact with the team structure of the economy. This is first-order for evaluating how different majors and fields contribute to aggregate innovation. Policy should be tilted towards building expertise that has large aggregate effects, and the policy must recognize the contribution that expertise makes to their productivity alone and to productivity in teams. This paper provides a benchmark for evaluating this issue from a general equilibrium framework.

D.4 Discussion on Policy Counterfactuals

In order to solve for a counterfactual scenario of an increase of the supply of type x given the existing distribution of skills, I need to solve a high dimensional nonlinear system of equations. However, this process can be simplified by an understanding of the Walrasian equilibrium and methods of tatonnement. The key element is to realize the excess demand function for each type is linked through the team formation equation as follows.

$$m_k = \exp\left(\frac{1}{T_k}\sum_{(x,\ell)\in k}\log\frac{m_k^{x,\ell}}{N_k^{x,\ell}} + V_k\right)$$

V_k represents the net value of the team as estimated in the previous section and is known. It has been identified in the previous equilibrium. m_k and $m_{x,\ell}$ are not observed in the counterfactual world. Given the identification of the model, there is a set of k equations for each team, and the market clearing for each type:

$$\sum_k N_k^{x,\ell} m_k = M_{x,\ell}$$

The overall process gets unwieldy. The following definition characterizes a vector of the mass of types alone:

[22] Jones and Williams (1998) discuss this misalignment and find that there is significant underinvestment in R&D.

$$\overrightarrow{m_0} = (m_{10}, ..., m_{\underline{k}})$$

And the excess demand equation for each type:

$$D^{\epsilon}_{x}(\overrightarrow{m_0}) = \sum m^{x}_{k}(\overrightarrow{m_0}) - M_x$$

Each excess demand function can be written out as a function of each type working alone, as each team equation is a function of the types working alone. The key condition is the following condition.

$$\begin{bmatrix} D^{\epsilon}_{1}(\overrightarrow{m_0}) = 0 \\ \vdots \\ D^{\epsilon}_{x}(\overrightarrow{m_0}) = 0 \\ \vdots \\ D^{\epsilon}_{X}(\overrightarrow{m_0}) = 0 \end{bmatrix}$$

The key result for counterfactuals is to find the tattonment equilibrium that satisfies these conditions. This enables the second quantitative exercise that explores changes in the supply of types. Instead of a high dimensional unwieldy equation, there are simply the same number of equations as the number of types.

D.5 Rank of Expertise

Here, I rank expertise by their contribution to aggregate innovation and note their partial equilibrium rank. The general equilibrium rank is based on increasing the supply of the given type by 0.01 and observing the corresponding patterns across teams. The partial equilibrium rank relies on Proposition 3. Table D.9 details the overall ranking, using data from 1991-2000:

Table D.9: A Rank of Expertise Across IPC3 Categories

Expertise Type	GE Rank	PE Rank	Proportion of Pop.
CHECKING-DEVICES	1	1	0.0048663
MEDICAL OR VETERINARY SCIENCE;	2	2	0.1254494
COMPUTING; CALCULATING; COUNTI	3	5	0.1081146
ELECTRIC COMMUNICATION TECHNIQ	4	6	0.0671355
SIGNALLING	5	8	0.0042086
BASIC ELECTRIC ELEMENTS	6	11	0.0856206
MUSICAL INSTRUMENTS; ACOUSTICS	7	14	0.0054672
CONTROLLING; REGULATING	8	13	0.005233
BIOCHEMISTRY; BEER; SPIRITS; W	9	7	0.0121443
EARTH OR ROCK DRILLING; MINING	10	15	0.0095658
COMBINATORIAL TECHNOLOGY [2006	11	9	0.0001055
INFORMATION STORAGE	12	18	0.0239818
OPTICS	13	10	0.0167979
NANOTECHNOLOGY [7]	14	22	0.0032794
ELECTRIC TECHNIQUES NOT OTHERW	15	25	0.0059509
PAPER-MAKING; PRODUCTION OF CE	16	31	0.0022594
DYES; PAINTS; POLISHES; NATURA	17	16	0.0059914
MEASURING; TESTING	18	24	0.0511502

GENERATION, CONVERSION, OR DIS	19	32	0.0129695
EDUCATING; CRYPTOGRAPHY; DISPL	20	34	0.0070125
BASIC ELECTRONIC CIRCUITRY	21	37	0.0192122
PRINTING; LINING MACHINES; TYP	22	26	0.0098581
VEHICLES IN GENERAL	23	52	0.0255578
CRYSTAL GROWTH [3]	24	40	0.0006626
GRINDING; POLISHING	25	48	0.0045533
LAYERED PRODUCTS	26	30	0.0043826
COATING METALLIC MATERIAL; COA	27	12	0.0049353
ORGANIC MACROMOLECULAR COMPOUN	28	17	0.0230767
ELECTROLYTIC OR ELECTROPHORETI	29	20	0.00171
ANIMAL OR VEGETABLE OILS, FATS	30	23	0.0028558
SPORTS; GAMES; AMUSEMENTS	31	51	0.0189342
BRAIDING; LACE-MAKING; KNITTIN	32	43	0.0013564
PHYSICAL OR CHEMICAL PROCESSES	33	44	0.0159891
FOOTWEAR	34	45	0.0020525
LIGHTING	35	46	0.0039711
DISPOSAL OF SOLID WASTE; RECLA	36	47	0.0006133
MACHINES OR ENGINES IN GENERAL	37	67	0.0051312

CEMENTS; CONCRETE; ARTIFICIAL	38	27	0.0028613
WORKING OF PLASTICS; WORKING O	39	55	0.0090339
COMBUSTION ENGINES; HOT-GAS OR	40	64	0.008281
GLASS; MINERAL OR SLAG WOOL	41	49	0.0022911
SEPARATING SOLIDS FROM SOLIDS;	42	59	0.0009044
INORGANIC CHEMISTRY	43	19	0.0022462
EXPLOSIVES; MATCHES	44	38	0.0006276
POSITIVE-DISPLACEMENT MACHINES	45	63	0.0044701
NATURAL OR MAN-MADE THREADS OR	46	42	0.0010456
PHOTOGRAPHY; CINEMATOGRAPHY; A	47	33	0.0138867
NUCLEAR PHYSICS; NUCLEAR ENGIN	48	41	0.0009646
PETROLEUM, GAS OR COKE INDUSTR	49	36	0.0049342
SEPARATION OF SOLID MATERIALS	50	53	0.000911
TREATMENT OF WATER, WASTE WATE	51	56	0.0041637
ORGANIC CHEMISTRY [2]	52	21	0.030416
TOBACCO; CIGARS; CIGARETTES; S	53	88	0.0007315
OPENING OR CLOSING BOTTLES, JA	54	72	0.0015863
RAILWAYS	55	82	0.0016563
HOROLOGY	56	99	0.0005717

STORING OR DISTRIBUTING GASES	57	54	0.0004174
SPRAYING OR ATOMISING IN GENER	58	60	0.0053599
AIRCRAFT; AVIATION; COSMONAUTI	59	81	0.0033549
SUGAR INDUSTRY [4]	60	4	0.0000672
COMBUSTION APPARATUS; COMBUSTI	61	84	0.0026019
FOODS OR FOODSTUFFS; THEIR TRE	62	50	0.0050688
POSITIVE-DISPLACEMENT MACHINES	63	61	0.0001329
REFRIGERATION OR COOLING; COMB	64	57	0.0048784
HEATING; RANGES; VENTILATING	65	85	0.0028558
METALLURGY; FERROUS OR NON-FER	66	69	0.0019474
CONVEYING; PACKING; STORING; H	67	86	0.0237935
HABERDASHERY; JEWELLERY	68	109	0.0014232
MAKING ARTICLES OF PAPER, CARD	69	62	0.0010686
LIFE-SAVING; FIRE-FIGHTING	70	79	0.0015578
AMMUNITION; BLASTING	71	65	0.0016049
LAND VEHICLES FOR TRAVELLING O	72	93	0.0065091
BOOKBINDING; ALBUMS; FILES; SP	73	80	0.0023841
LOCKS; KEYS; WINDOW OR DOOR FI	74	92	0.0041538
CENTRIFUGAL APPARATUS OR MACHI	75	58	0.0004109

MACHINE TOOLS; METAL-WORKING N	76	76	0.007829
ENGINEERING ELEMENTS OR UNITS;	77	87	0.0204665
WORKING CEMENT, CLAY, OR STONE	78	77	0.0005291
DOORS, WINDOWS, SHUTTERS, OR R	79	91	0.0026742
GENERATING OR TRANSMITTING MEC	80	28	0.0001745
WEAPONS	81	83	0.0040893
AGRICULTURE; FORESTRY; ANIMAL	82	102	0.019893
MICROSTRUCTURAL TECHNOLOGY [7]	83	35	0.0000234
TREATMENT OF TEXTILES OR THE L	84	66	0.0022014
BRUSHWARE	85	70	0.0011934
HYDRAULIC ENGINEERING; FOUNDAT	86	96	0.0031185
BUILDING	87	95	0.0086575
CLEANING	88	68	0.0015983
HEAT EXCHANGE IN GENERAL	89	73	0.0010949
DRYING	90	74	0.00061
FURNITURE; DOMESTIC ARTICLES O	91	94	0.0200035
WEARING APPAREL	92	97	0.0029478
WORKING OR PRESERVING WOOD OR	93	98	0.0012459
CONSTRUCTION OF ROADS, RAILWAY	94	100	0.0023042

WATER SUPPLY; SEWERAGE	95	116	0.0017877
DECORATIVE ARTS	96	107	0.0008333
FERTILISERS; MANUFACTURE THERE	97	75	0.0004119
CASTING; POWDER METALLURGY	98	78	0.0020875
HAND TOOLS; PORTABLE POWER-DRI	99	111	0.0058863
METALLURGY OF IRON	100	71	0.0008432
WEAVING	101	103	0.0003222
STEAM GENERATION	102	39	0.0002565
YARNS; MECHANICAL FINISHING OF	103	114	0.0001843
BAKING; EQUIPMENT FOR MAKING O	104	90	0.0006965
HAND CUTTING TOOLS; CUTTING; S	105	104	0.0019912
HOISTING; LIFTING; HAULING	106	108	0.0021215
FLUID-PRESSURE ACTUATORS; HYDR	107	101	0.0009581
CRUSHING, PULVERISING, OR DISI	108	118	0.0012142
HEADWEAR	109	119	0.0009187
HAND OR TRAVELLING ARTICLES	110	117	0.0046036
MECHANICAL METAL-WORKING WITHO	111	106	0.0036701
BUTCHERING; MEAT TREATMENT; PR	112	120	0.0008246
MACHINES OR ENGINES FOR LIQUID	113	89	0.000169

SEWING; EMBROIDERING; TUFTING	114	105	0.0005597
SKINS; HIDES; PELTS; LEATHER	115	29	0.0000376
WRITING OR DRAWING IMPLEMENTS;	116	112	0.0007217
PRESSES	117	115	0.0004371
SHIPS OR OTHER WATERBORNE VESS	118	121	0.0038879
FURNACES; KILNS; OVENS; RETORT	119	113	0.0003299
SADDLERY; UPHOLSTERY	120	122	0.0002007
ROPES; CABLES OTHER THAN ELECT	121	110	0.0000464
ROPES; CABLES OTHER THAN ELECT	122	3	0.0000147

Notes: This table lists the skills at the IPC3 level (Expertise Types) and evaluates their rank in general equilibrium (e.g. from a supply shock to the economy and corresponding change in aggregate innovation) and private rank (e.g., partial equilibrium, the marginal value of entering the market as the expertise type. Source: USPTO and author calculations.