NBER WORKING PAPER SERIES

WHY ARE SOME REGIONS MORE INNOVATIVE THAN OTHERS? THE ROLE OF FIRM SIZE DIVERSITY

Ajay K. Agrawal
Iain M. Cockburn
Alberto Galasso
Alexander Oettl

Working Paper 17793
http://www.nber.org/papers/w17793

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2012

We thank Ashish Arora, Joshua Gans, Avi Goldfarb, Nico Lacetera, Matt Mitchell, Mark Schankerman, Will Strange and seminar participants at the University of Toronto, the London School of Economics, the London Business School, the Queen's Conference on Innovation and Entrepreneurship, the REER conference at Georgia Tech, and the EARIE conference in Stockholm. We gratefully acknowledge generous financial support from the Martin Prosperity Institute, the Centre for Innovation and Entrepreneurship at the Rotman School of Management, and the Social Sciences and Humanities Research Council of Canada for funding this research. All errors and omissions are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Why are Some Regions More Innovative than Others? The Role of Firm Size Diversity
Ajay K. Agrawal, Iain M. Cockburn, Alberto Galasso, and Alexander Oettl
NBER Working Paper No. 17793
January 2012
JEL No. O31,R11

ABSTRACT

Large labs may spawn spin-outs caused by innovations deemed unrelated to the firm's overall business. Small labs generate demand for specialized services that lower entry costs for others. We develop a theoretical framework to study the interplay of these two localized externalities and their impact on regional innovation. We examine MSA-level patent data during the period 1975-2000 and find that innovation output is higher where large and small labs coexist. The finding is robust to across-region as well as within-region analysis, IV analysis, and the effect is stronger in certain subsamples consistent with our explanation but not the plausible alternatives.

Ajay K. Agrawal
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, ON M5S 3E6
CANADA
and NBER
ajay.agrawal@rotman.utoronto.ca

Iain M. Cockburn
School of Management
Boston University
595 Commonwealth Ave
Boston, MA 02215
and NBER
cockburn@bu.edu

Alberto Galasso
Rotman School of Management
University of Toronto
105 St. George Street
Toronto, ON
CANADA M5S 3E6
Alberto.Galasso@Rotman.Utoronto.Ca

Alexander Oettl
College of Management
Georgia Institute of Technology
800 West Peachtree Street, NW
Atlanta, GA 30308
alexander.oettl@mgt.gatech.edu

1 Introduction

A striking feature of economic geography is the large variation in innovation productivity across regions. Silicon Valley and Boston are popular examples of regions that are significantly more productive than others in terms of innovation. In Figure 1, we illustrate a broader cross section of such variation using patent data on US computers and communications. Even Metropolitan Statistical Areas (MSAs) of a similar size in terms of the number of local inventors often differ substantially in terms of their innovation productivity (number of citation-weighted patented inventions per inventor). For example, Rochester and Portland had a similar number of innovators working in the computer and communications industry in 1995, but Portland inventors generated almost double the number of citation-weighted patents.

Regional productivity disparities have led to a variety of policies focused on enhancing local innovation. Such initiatives often focus either on encouraging entrepreneurship (e.g., San Diego, CA, New York, NY, and St. Louis, MO) or on attracting large corporate labs (e.g., Flint, MI, Greenville, SC, and Shelby, AL).[1] We argue that effective regional innovation policymaking requires an understanding of how the structure of local R&D manpower is related to innovation productivity.

In this paper, we study how local innovation is affected by the organization of R&D manpower in that region. For over six decades, since Schumpeter (1942), innovation scholars have tried to understand the relationship between product market industry structure and innovation (Geroski (1990), Grossman and Helpman (1991), Aghion et al. (2005)). Other research has focused on the relationship between innovation and regional industrial diversity, for example, comparing innovative output from cities focused primarily on one industry (e.g., automobiles) with industrially diverse cities (e.g., electronics, chemicals, and textiles) (see Jacobs (1969), Glaeser et al. (1992), and Feldman and Audretsch (1999)). Despite this extensive literature, the effect of R&D labor organization on local innovation has so far attracted little empirical and theoretical attention. We fill

[1]The 2008 Brookings Institute's Blueprint for American Prosperity offers a comprehensive overview of such regional initiatives. For example, in San Diego, the CONNECT program has helped the development of more than 2,000 small firms in the hi-tech and bio-tech sectors since 1985. New York recently launched the NYC High-Tech CONNECT program modeled on San Diego's CONNECT. Similarly, St. Louis implemented a number of policies to promote regional entrepreneurship. Flint, Greenville, and Shelby focused instead on attracting large firms - GM, BMW, and Mercedes, respectively.

this gap by combining insights from urban economics and entrepreneurship.[2]

We develop a simple theoretical framework to study the impact of R&D labor organization on innovation. The model relies on three main assumptions. First, following Schumpeter (1942), we assume that large labs have an advantage in idea production.[3] Second, following Cassiman and Ueda (2006), we assume that large labs only commercialize innovations that "fit" with their established research activities. However, if potentially profitable, then spin-outs may commercialize "misfit" inventions that do not fit with the assets, mandate, or strategy of the parent firm.[4] Third, we assume that the cost of spin-out formation is reduced if a large number of small labs is present. This assumption follows Vernon (1960) and Chinitz (1961), who argue that an increasing number of small firms "thicken" local markets for ancillary services and thus reduce entry costs.

We develop a model that shows how the manner in which regional R&D manpower is organized may have an impact on local innovation. The main testable prediction is that innovation productivity is greater in "diverse" MSAs, which we define as MSAs where at least one large lab and numerous small labs coexist compared to MSAs of a similar size but with only large or small labs. Because of the large lab advantage in idea production, an increase in concentration of large labs (keeping constant the number of small labs) increases spin-out formation. Because of small firm market thickness externalities, an increase in the number of small labs (keeping constant the number and size of large labs) renders a spin-out more profitable. Under certain conditions, this trade-off implies that spin-out formation is maximized when at least one large lab and numerous small labs

[2]See Cohen (2010) for an excellent survey of the the 'neo-Schumpeterian' empirical literature. Our paper also contributes to the literature on spin-out formation. While this literature has explored the impact of parent firm characteristics on spin-out performance (Franco and Filson (2006)) and contrasted spin-outs to other entrants (Chatterji (2009)), our paper is to our knowledge the first to examine the impact of regional R&D manpower organization on local spin-out formation.

[3]This typically arises when the lab can spread R&D fixed costs over a larger number of innovations (see Cohen and Klepper (1996) for a micro-foundation). Empirical evidence of such an advantage is provided in Klette (1996), Henderson and Cockburn (1996), and Cockburn and Henderson (2001). Alternatively, scale advantages may arise from division of labor efficiencies (Arora and Gambardella (1994)) or human capital complementarities (Jones (2008)).

[4]Prominent examples of such spin-outs include: Intel, founded by Andy Grove, Bob Noyce and others to make a product that Fairchild was unwilling to make; Lotus Development founded by Mitch Kapor that left Digital Equipment Corporation; and FreeMarket, founded by a General Electric (GE) engineer after GE rejected his initial proposal. In 2002 the Wall Street Journal reported that in 2001 GE's researchers suggested more than 2,000 new products but only five proposals were accepted for product development (see Cassiman and Ueda (2006) for additional examples).

are present. Because spin-outs allow innovators to commercialize inventions that would otherwise be abandoned since they are not a good fit with their employer's research activities, the number of commercialized inventions also increases when both types of labs are present.

We test the empirical predictions of the model using a 26-year panel dataset at the MSA-technology-year level.[5] The data strongly support the theoretical model. First, focusing on across-region variation, we find that locations with firm size diversity in 1995 have an average 46% innovation premium five years later compared to those without. Recognizing the potential for omitted variable bias, we turn to within-region-technology class variation over time (1975-2000) and use MSA-technology fixed effects in our main specification. We find that in periods where at least one large lab and numerous small labs co-exist, MSAs experience a 17% increase in citation-weighted patent counts. To further address endogeneity, we show this result is robust to using lagged income tax rates, which vary at the state level, as an instrument for regional firm size diversity. High income tax rates induce entry and thus may generate subsequent regional firm size diversity. Finally, we show that the main result is robust to disaggregating the firm size diversity measure into separate measures for large and small labs, to focusing on a smaller sample of just large MSAs, to focusing on a smaller sample that drops California MSAs, to applying different measures of diversity with different cutoffs for large and small firms, and to an alternate measure of diversity that uses County Business Patterns Census data rather than patent data.

Next, we turn to examining whether the data are consistent with the *mechanism* we focus on in our model that links firm size diversity to regional innovation. First, we show that diversity is associated with a 32% increase in the probability of spin-out formation. Second, since the main channel through which diversity increases innovation in our model is spin-out formation, we expect that any barrier to spin-out formation will

[5] A number of case studies also provide support for our theory. For example, consider Portland, OR versus Rochester, NY (lack of small firms) and Atlanta, GA versus Seattle, WA (lack of large firms) in 1995. In terms of Portland and Rochester, the number of inventors patenting in the "computers and communications" technology class is very similar in the two cities (roughly 1,000 inventors). Nonetheless, Portland outperforms Rochester, obtaining almost 50% more patents and about twice the number of citation-weighted patents than Rochester. While both cities register a similar presence of large labs, the number of small labs is substantially different: Portland has more than five times the number of small labs as Rochester. On the other hand, in the "chemicals" technology class, Seattle and Atlanta have a similar number of small labs (38 and 36, respectively) and also a similar number of overall inventors (457 and 484, respectively), but only Atlanta has a large lab (Kimberly Clark). The difference in innovation output: Atlanta has 37% more citation-weighted patents.

reduce the beneficial effect of firm size diversity. We show that the effect of firm size diversity on innovation is indeed reduced by the presence of strong non-compete laws. Third, since spin-out formation in our model is predicated on ideas produced by large labs that are subsequently deemed unrelated, we expect that regions with large labs that maintain a narrower focus and thus produce more "misfit ideas" will benefit more from firm size diversity. We show that the effect of firm size diversity on innovation is indeed higher in regions with more narrowly focused large labs.

We exploit these mechanism-related findings to assess the validity of our theory against competing explanations for the positive association between regional firm size diversity and innovation. Several alternative theories are consistent with a subset of the correlations we report but none with the whole set.

We organize the paper as follows. We present a model in Section 2 and describe our data in Section 3. We introduce the empirical framework in Section 4, explain our results in Section 5, and conclude in Section 6. We include proofs of the theoretical results as well as additional empirical results in the Appendices.

2 Model

Consider an MSA with T scientists divided into J large labs and N small labs. A small lab employs only one scientist and cannot commercialize multiple innovations. Large labs employ at least two scientists and have the capability to commercialize more than one innovation. Moreover, each large lab j has an existing research activity that generates profits λ_j. We indicate as $S_j > 1$ the number of scientists working in large lab j and denote with $S_L = \sum_{i=1}^{J} S_j$ the total number of scientists working in large labs. Notice that $T = S_L + N$.

We assume there are economies of scale in innovation: each small lab discovers an innovation with probability $1/\Delta$, and each scientist in a large lab of size S_j discovers an innovation with probability S_j/Δ with $\Delta \geq T$.

If commercialized in a small lab, an innovation generates profits equal to π with $1 \geq \pi > 0$. For a large lab, the profits from the commercialization of an innovation discovered internally are equal to $\pi + a$. The parameter $a \in R$ captures the "fit" of the innovation with the existing research activity of the lab (Cassiman and Ueda, 2006). A positive value for a may arise because of scale or scope economies. Conversely, a negative value for a indicates the existence of coordination costs between the innovation and the

existing research activity. Alternatively, $a < 0$ may indicate the opportunity cost of deploying resources from the existing research activity.[6] Notice that an innovation is profitable for a large lab j if the profits from commercialization $(\pi + a + \lambda_j)$ exceed those without commercialization (λ_j), i.e., when $\pi + a > 0$.

A scientist working in a large lab can implement the discovered innovation outside the lab by opening a new small lab (spin-out). The cost of opening a new lab is equal to $k(N)$ with $k(0) > 1$, $k'(N) < 0$, and $\lim_{N\to\infty} k(N) = 0$. The entry cost is decreasing in N because the presence of other small labs generates a positive externality.[7]

The timing of the game is as follows. In the first period, innovations are discovered. In the second period, large lab scientists negotiate with their labs about the destiny of their innovations. The possible strategies are: (i) internal commercialization (with profits $\pi + a + \lambda_j$), (ii) spin-out formation (with profits $\pi - k(N)$), and (iii) non-commercialization (with profits λ_j). In the third period, spin-outs are formed and ideas are commercialized.

Commercializing an innovation (internally or through a spin-out) may require the development of subsequent technologies necessary to extract profits from the idea. If this occurs, commercialized innovations generate new research trajectories. Conversely, the non-commercialization strategy produces *orphan ideas* that do not generate follow-on research.[8]

To solve the bargaining problem at period 2, we assume that the lab makes a take-it-or-leave-it offer to the scientist.

Proposition 1 *Established small labs commercialize internal innovations with* $\pi \geq 0$. *Large labs commercialize internal innovations with* $\pi + a \geq 0$ *and* $a \geq -k(N)$. *A spin-out is formed if* $\pi \geq k(N)$ *and* $a \leq -k(N)$.

Figure 2 summarizes the destiny of an innovation discovered in a large lab. When

[6] See Bresnahan et al. (2011) for evidence of scope diseconomies in the computer industry.

[7] A possible micro-foundation of this cost can be provided building on the model of Helsley and Strange (2002) where a large number of small labs generates a dense network of input suppliers that facilitates spin-out formation. An alternative interpretation for $k(N)$ is that in the presence of small firms, the MSA develops a culture of entrepreneurship that induces employees of large labs to leave their employer and start their own firms (Glaeser and Kerr (2009)).

[8] Empirically, this implies that greater commercialization may not be associated with greater patenting (because orphan ideas may also be patented) but is likely to be associated with greater citation-weighted patent counts (because orphan ideas do not generate new research trajectories).

the cost of establishing a spin-out is large, scientists do not find it profitable to form a spin-out and thus the large lab commercializes the innovations if the fit with the established business is high enough. When the cost of forming a spin-out is low ($k \leq \pi$), then inventors form spin-outs to commercialize innovations that are not a good fit with the existing research activity.

We assume that a is distributed with a continuous and differentiable cumulative distribution $F(a)$. Under this assumption, a spin-out is formed with a positive probability as long as $\pi \geq k(N)$ that is satisfied if $N \geq \underline{N}$ where $\underline{N}$ is defined as $\pi - k(\underline{N}) = 0$. Therefore, an innovation of a large lab scientist will generate a spin-out with probability:

$$\Pr(\textit{spin-out}) = \begin{cases} F\left(-k(N)\right) & \text{if} \quad N \geq \underline{N} \\ \\ 0 & \text{if} \quad N < \underline{N} \end{cases}.$$

Similarly, an innovation discovered in a large lab is commercialized (either internally or through spin-outs) with probability:

$$\Pr(\textit{Comm. from Large}) = \begin{cases} 1 & \text{if} \quad N \geq \underline{N} \\ \\ 1 - F(-\pi) & \text{if} \quad N < \underline{N} \end{cases}$$

whereas each scientist in a small lab discovers an innovation with probability $1/\Delta$. Because in large lab j each scientist discovers an innovation with probability S_j/Δ, the expected number of innovations in lab j is S_j^2/Δ. Exploiting these results, we obtain the following proposition.

Proposition 2 *The expected number of spin-outs, NS, is maximized in the presence of one large lab and $N \geq \underline{N}$. If T is large enough, then the expected number of commercialized innovations, NI, is maximized in the presence of one large lab and $\underline{N}$ small labs.*

Proposition 2 indicates that the presence of large and small labs are complementary forces affecting both innovation and spin-out formation. Intuitively, a greater concentration among large labs increases the number of ideas generated and the expected number of spin-outs. A larger pool of existing small labs facilitates the formation of spin-outs and thus allow ideas that are not a "good fit" for a large lab to be commercialized. There

is a positive interaction between the two effects. A large number of small labs has no effect on spin-out formation in the absence of large labs. At the same time, an increase in large lab concentration generates a greater number of spin-outs in the presence of small lab externalities.[9]

Let us label an MSA as *diverse* if there are $N \geq \underline{N}$ small labs and only one large lab. By showing that spin-out formation and the number of commercialized innovations are maximized with diverse MSA configurations, Proposition 2 implies that any non-diverse MSA configuration is dominated by at least one diverse configuration in terms of spin-out formation or innovation.

Building on these results, in the Appendix we show that the distinction between diverse and non-diverse MSAs is even starker. For an arbitrarily diverse MSA with $N \geq \underline{N}$ small labs and only one large lab, we show that if the MSA is large enough, there is no non-diverse MSA configuration that generates more innovation and spin-outs.[10]

3 Data

We focus on two units of analysis. First, we study cross-region variation and use MSA-technology class as our unit of analysis (e.g., Rochester, NY - chemicals). Then, we turn our attention to within-region variation and use MSA-class-year as our unit of analysis.

In constructing our sample, we begin with the set of 268 MSAs defined in 1999 by the US Office of Management and Budget[11] and the set of six one-digit technology

[9]For simplicity, we do not consider: (i) cannibalization and complementarity effects between the innovations developed by the spin-outs and those commercialized by the large lab and (ii) imperfect IP protection. In a previous draft of the paper, we showed it possible to obtain similar results in a model with cannibalization, complementarity, and imperfect IP protection as in Gans and Stern (2000) and in a model where spin-outs are formed because of disagreements between large lab scientists as in Klepper and Thompson (2010). Intuitively, as long as the IP regime allows spin-outs to be profitable, the presence of local externalities encourages spin-out formation (formal proofs available upon request).

[10]In the previous analysis, we refer to an MSA as "diverse" if there are $N \geq \underline{N}$ small labs and only one large lab. In the Appendix, we relax this definition and consider MSAs where multiple large labs are present. We show that in this environment, it is possible to characterize a size threshold $\underline{S}$ such that if at least one large lab exceeds this threshold, then the MSA outperforms all non-diverse MSAs with $N < \underline{N}$. This result provides a micro-foundation of our main empirical measure of firm size diversity that is based on two cut-offs: one for the number of small labs, $\underline{N}$, and another for the size of large labs, $\underline{S}$.

[11]`http://www.census.gov/population/estimates/metro-city/99mfips.txt`

classes described in Hall et al. (2001).[12] This generates 1608 MSA-class observations. We then construct our panel dataset, which includes 26 years (1975-2000) and thus contains 41,808 MSA-class-year observations.

We measure innovative activity, our main dependent variable, using a citation-weighted count of US patents:

Weighted Patents$_{jkt+5}$: the forward citation weighted sum of distinct patents with primary technology classification k and application year $t+5$ where at least one inventor is located in MSA j.

We use inventor address information to assign a patent to an MSA, exploiting the US National Geospatial-Intelligence Agency dataset to match cities and townships to counties and ultimately MSAs. If a patent has at least one inventor from a particular MSA, then we increment the counter for that MSA by one. Thus, a patent with three inventors located in three different MSAs increments the patent counter for each of those MSAs by one. However, if all three inventors are located in the same MSA, then the counter for that MSA is only incremented by one.[13]

We construct this measure using all patents applied for (and subsequently granted) between 1975 and 2000, with at least one inventor with a US address. We exclude patents that cannot be attributed to an MSA (due to incomplete address information or a location outside an MSA) and patents assigned to universities and hospitals. The US Patent and Trademark Office is the original source of our patent data. We complement these data with classification data from the NBER (technology classification, assignee name).

Patent citations identify prior knowledge upon which a patent builds, and prior literature (starting with Pakes and Griliches (1980)) has often employed the number of forward-citations received by a patent as an indirect measure of patent value.

We also consider an unweighted patent count as an additional innovation metric:

Patents$_{jkt+5}$: the number of distinct patents with primary technology classification k and application year $t+5$ where at least one inventor is located in MSA j.

Our second dependent variable is a measure of spin-out formation, which we use in the latter part of our analysis when we turn our attention to the mechanism through

[12]This classification scheme includes: chemicals, computers and communications, drugs and medical, electrical and electronic, mechanical, and other.

[13]Results are robust to constructing this variable using first inventor data only.

which firm size diversity influences innovation output. We define a spin-out as a particular type of lab. We define a lab in MSA j, technology class k, in year t as follows. First, we take all assignees that, within a five-year window (year t and the four preceding years), applied for at least one (eventually granted) patent in technology class k. Second, using this list of assignees, we identify labs in MSA-class jk if there are at least three different inventors located in MSA j who are named in class k patents of that assignee during the five-year window. Thus, if a firm has operations (i.e., R&D labs or facilities) in n different MSA-classes, our procedure will treat it as n distinct entities. In any given year, the number of patents attributed to such "labs" are thus computed by aggregating by (standardized) assignee name, location, and application date.

We define a spin-out as a new lab in MSA-class jk if among the patents applied for during the first year of lab activity we identify at least one inventor who previously patented in one of the large labs in MSA jk. We construct the following variable to measure spin-out formation:

log(Spin-Outs$_{jkt}$): logarithm of one plus the count of spin-outs in year t MSA-class jk.

Finally, our main explanatory variable is a measure of diversity. We construct this variable by first identifying the distribution of lab sizes in each technology class k and year t. By construction, lab size has a lower bound of three (each lab has at least three inventors). Across the various class-years, the median size is about five inventors, the 75th percentile is about nine inventors, and the 97th percentile is roughly 54 inventors. We use this distribution to define large and small labs. A *large* lab is a lab where the number of inventors is above the 97th percentile in the technology class-year distribution. We define a lab as *small* if the number of inventors is below the 75th percentile. We arbitrarily choose the size thresholds for large and small labs (97th and 75th percentile), but we perform robustness checks in the Appendix.

Diverse$_{jkt}$: dummy variable equal to one if in year t MSA-class jk has at least one active large lab and at least 139 active small labs.

The number of small labs (139) corresponds to the 99th percentile of the distribution of the number of small labs across the entire sample. Also, for this cut-off, we perform a number of robustness checks in the Appendix. This measure intends to capture MSAs where both large labs and many small labs coexist. Notice that the diverse dummy is distinct from a count of the number of labs because it considers lab sizes. It

is also distinct from traditional concentration measures (Herfindahl or share of top four labs) because intermediate values of these measures may emerge both with and without diversity.[14]

Table 1 presents descriptive statistics for the main variables. The average MSA-class has approximately 91 inventors and eight labs. On average, the inventors of an MSA-class apply for about 28 patents per year, and these patents receive roughly 466 forward citations. The distributions of these variables are highly skewed. The median MSA-class in our sample has only one lab and five inventors who apply for three patents per year and receive 33 cites.

The *diverse* dummy equals one for about 1% of the sample. The small fraction of diverse observations is mostly due to the fact that a large number of MSA-classes do not have enough inventors to display a positive diversity measure. In Table 2 we focus on MSA-class-years that are "at risk" of becoming diverse.[15] The *diverse* dummy equals one for roughly 16% of this smaller sample. On average, these MSA-classes have approximately 105 labs and 1348 inventors, who apply for about 381 patents per year that receive roughly 6681 forward citations.

Table 3 illustrates the variation in diversity status that will be exploited in our empirical analysis. We observe a switch in diversity status for 36 MSA-classes. The first MSA-class to become diverse is New York (in the chemical technology area) in 1976, although switches occur throughout the entire sample period. There are 11 unique MSAs experiencing a switch in at least one technology class, and all technology classes experience at least four switches.

4 Methodology

Our main econometric model focuses on the relationship between count-based measures of innovative activity Y_{jkt+5} in MSA-class jk in period $t+5$ and the indicator for firm

[14]To see this, consider the following example. MSA A has four identical labs, each employing a quarter of the local inventors. MSA B has one lab employing half of the local inventors as well as a very large number of small labs. In MSA A, the Herfindahl index is equal to 1/4 and in MSA B the Herfindahl index is also (approximately) 1/4. However, MSA B has large firm–small firm coexistence, while MSA A does not.

[15]The lab size distribution is bounded below by three (by definition each lab has at least three inventors), and across the various classes-years the 97th percentile is roughly 54 inventors. Therefore, our constructed cut-off is equal to 3x139+54=471 inventors.

size diversity $diverse_{jkt}$ in MSA-class jk in period t. We typically model the conditional expectation of innovative activity as:

$$E[Y_{jkt+5}] = \exp(\alpha diverse_{jkt} + x_{jkt}\beta + \gamma_{jk} + \lambda_{tk}) \tag{1}$$

where x_{jkt} is a vector of control variables, γ_{jk} is an MSA-class specific idiosyncratic effect, and λ_{tk} is a vector of technology class time-period effects. Notice that all the dependent variables are lagged five periods to account for simultaneity concerns.

Equation (1) uses the log-link formulation due to the non-negative and highly skewed nature of our count-based dependent variables. Following Wooldridge (1999), we adopt the Poisson quasi maximum-likelihood estimation that yields consistent estimates as long as the conditional mean is correctly specified. We cluster the standard errors to allow for arbitrary heteroskedasticity and autocorrelation.

When x in Equation (1) includes measures of the number of inventors working in the MSA-class, α indicates whether MSA-classes with a diverse configuration receive more citation-weighted patents per inventor; therefore, it is a measure of MSA-class productivity.

5 Results

Firm Size Diversity and Innovation

Across–Region Variation

Table 4 contains our first set of results, which show a robust positive association between firm size diversity and innovation in cross-section regressions. We estimate all models in Table 4 using Poisson, with robust standard errors clustered at the MSA-class level to account for over-dispersion. In Columns (1) to (3), the dependent variable is the citation-weighted patent count or, equivalently, the total forward citation count for issued patents applied for by all inventors in the MSA-class in year $t + 5$.

Column (1), focusing on year 1995, shows a large positive correlation between size diversity and innovation. In Column (2), we include a control for the number of inventors in the MSA-class. The positive coefficient on *diverse* now indicates that diverse MSAs obtain more innovation per inventor. In Column (3), we show that the correlation is similar when we control for technology effects. Exponentiation of the coefficient in

Column (3) implies that diverse MSA-classes have a 46% innovation premium over non-diverse MSA-classes. In Column (4), we show a similar positive correlation measuring innovation with un-weighted patent counts.

Within–Region Variation

Focusing on year 1995, the cross-section regressions reported in Table 4 indicate a positive association between diversity and innovation. Column (1) of Table 5 confirms this result in a pooled cross section that exploits the entire sample period. While this correlation is consistent with our theory, the result may be due to unobserved MSA-class heterogeneity that is correlated with diversity and innovation. Moreover, the previous regressions include MSA-classes that do not have enough inventors to display a positive diversity measure. To address these concerns, in Column (2) we move to a fixed-effects Poisson estimator (Hausman et al. (1984)) with MSA-class fixed effects, year effects, and technology class time trends and drop all the MSA-classes that are too small to be diverse. This specification isolates the within MSA-class co-variation of diversity and innovation. The estimated coefficient implies an increase in the citation-weighted patent count of about 17% in periods where MSAs are diverse.

Column (3) shows that the qualitative and quantitative results are robust to introducing additional controls including the Herfindahl concentration index for the labs in the MSA-class and the number of active labs. This confirms that the diversity measure is not simply capturing lab concentration or fragmentation. In Columns (4) and (5), we show that results are similar in the larger dataset that includes small MSA-classes. The magnitude and statistical significance of the coefficients are similar with longer or shorter lag structures.[16] Finally, in Column (6), we look at unweighted patent counts. We still find a positive and significant correlation between diversity and innovation, but the magnitude of the effect drops to 6.5%, suggesting that diversity has a greater impact on the quality-adjusted measure than on the number of patents.[17] Overall, the results

[16]The coefficient on diversity equals 0.180 (p-value =0.013) with a seven-period lag, 0.119 (p-value =0.01) with a two-period lag, and 0.085 (p-value =0.032) with a one-period lag. The correlation between diversity and contemporaneous cites is positive but not statistically significant (equal to 0.048 p-value =0.12).

[17]In our model, for simplicity, we assume that all innovations have the same quality and differ only in their fit with the existing research activity. Introducing quality heterogeneity would still generate the prediction that diversity is associated with an increase both in quality (total number of cites) and in the number of patents. Whether the effect is stronger on quality-adjusted patents than on the number

in Table 5 document a strong positive association between diversity and innovation.[18]

The diversity dummy is a natural empirical construct to assess the correlation between innovation and diversity predicted by our theoretical model. In Appendix 2, we show this correlation is present in regressions with both direct effects and interaction (Table A2), confirming that an innovation premium is associated with the co-existence of large and small labs and that the presence of only large or small labs does not drive these results.

Instrumental Variable: Lagged Income Tax Rates

In Table 5, we introduce MSA-class fixed effects to control for time-invariant heterogeneity affecting both local R&D market structure and innovation. To identify the causal effect of diversity on innovation, we also need to address the potential bias arising from the correlation between time variant unobservable heterogeneity and firm size diversity. This correlation can arise in a variety of ways. A positive shock in the value of the technologies produced in the MSA-class may lead to an increase both in the entry of small firms and in the likelihood of innovation. Similarly, capital market shocks (e.g., the dot com boom in the late 1990s) may facilitate expansion of large firms and the entry of new firms as well as increase the overall investment in R&D. Human capital shocks (e.g., an increase in the H1B visas cap) may also lead to an inflow of scientists in the MSA that may cause a change in both lab structures and innovation. Similarly, local productivity shocks due to availability of information and communication technologies (Agrawal and Goldfarb (2008); Bloom et al. (2012)) may impact both the organization of R&D labor and innovation. Finally, successful innovation may lead empire-building CEOs to engage in takeovers and eliminate small firms from the MSA-class.

To address these endogeneity concerns, we need an instrumental variable that affects firm size diversity but does not directly affect innovation. We exploit variation in income tax rates over the sample period as an instrument for diversity. Many papers have documented a positive relationship between income tax rates and self-employment (among others, see Long (1982), Bruce (2000), Gentry and Hubbard (2004), and Cullen

of patents will depend on the correlation between the quality and the fit parameters.

[18]We obtain similar positive correlations: (i) in panel regressions in which we use the 36 two-digit categories defined by Hall, Jaffe, and Trajtenberg (2001) as the level of technology disaggregation and (ii) in regressions in which the dependent variable does not include cites obtained by large labs.

and Gordon (2007)). A common explanation for this finding is that self-employment offers tax-sheltering opportunities through the deduction of certain types of expenses (Cullen and Gordon (2007)).

If high income taxes induce some scientists to establish their own small labs, then we should expect income taxes to be positively correlated with diversity. Additionally, as long as states do not change income taxes in response to innovation shocks, we should expect variation in income tax rates to be uncorrelated with unobservable heterogeneity affecting innovation. Nonetheless, a possible problem with this instrument is that high income tax rates may also affect spin-out formation, thus complicating the interpretation of the estimates. To address this concern, we control for both contemporaneous and lagged (three years) income tax rates exploiting the lag tax rates as an instrument. Intuitively, we expect income tax rates in year $t-3$ to be correlated with market structure (diversity) in year t because of their effect on past entry decisions. At the same time, because we control for current income taxes, we expect lagged income taxes to be uncorrelated with the current decision to form a spin-out and with innovation in year $t+5$.

Following Galasso et al. (2011), we obtain information on state and federal income and capital gain taxes from the NBER Taxsim database described in Feenberg and Coutts (1993).[19] Our main tax variable is the combined (state plus federal) tax rate that a representative household faces in a specific MSA in a given year. Appendix Table A1 illustrates the variation of income tax rates averaged across states for five-year time periods. There is a substantial decline in tax rates in the late 1980s and an increase in the early 1990s. The table also indicates variation across US states with differences between the lowest and highest rates of 7 to 28 percentage points.

We present estimates of the instrumental variable regressions in Table 6. Column (1) reports the baseline OLS estimates with diversity non-instrumented. The coefficient on the diversity dummy implies an 18% increase in innovation that is very similar to the effect estimated with the baseline Poisson model. Column (2) presents coefficients of the first stage regression indicating that lagged income taxes are positively associated with diversity. The estimated impact of taxes on diversity is large: a one standard deviation increase in lagged income tax rates increases the probability of being diverse by approximately 25%. The regression shows that current income taxes are not correlated with

[19]This data set contains income tax rates by year and state for an additional $1,000 of income for a representative household (with $500,000 of wage income split evenly between husband and wife).

diversity. In unreported regressions, we also find no statistically significant correlation between spin-out formation and lagged income taxes (p-value =0.34). Together, these results suggest that lagged income taxes may serve as a valid instrument for diversity. Column (3) reports the second stage estimates with the diversity dummy instrumented by lagged income taxes. Qualitatively, we find that diversity is still positively associated with innovation. The coefficient of diversity is more than three times greater than the one in the OLS estimate, but the standard error is also larger and the two confidence intervals overlap.[20]

While most of the previous literature studying the impact of taxes on entrepreneurship focuses only on income tax rates, it is plausible that the effect of income taxes on entry depends on non-income taxes that new firms are required to pay. To this end, in Column (4), we control for capital gains tax rates. Because reductions in capital gain taxes may be associated with other pro-business policies that affect innovation, we include this variable on both stages of the IV regressions. In the (unreported) first stage, we notice that capital gains taxes have a negative impact on diversity. This is intuitive because high capital gain taxes reduce the profits that can be made by selling firm assets and shares and therefore reduce self-employment incentives. Moreover, high capital gain taxes may reduce the supply of venture capital that in turn will reduce entry of new firms. Column (4) shows that the second stage regressions are robust to the inclusion of capital gains taxes. The coefficient on diversity is still about three times larger than the one in the OLS estimates, suggesting that endogeneity generates a downward bias. Results are similar if we introduce additional controls, such as corporate taxes, the number of labs in the MSA-class, and the Herfindhal index.

Mechanism

Firm Size Diversity and Spin-Out Formation

The previous analysis illustrates a positive and robust effect of firm size diversity on innovation. Our theoretical framework indicates that the main channel through which diversity affects innovation is spin-out formation. This is consistent with previous literature indicating that spin-out innovation is superior to those of other entrants (Agarwal

[20]We obtain very similar estimates with a five-year lag for income taxes. Similarly, including both taxes in year t-3 and taxes in year t-5 as instruments leads to similar results.

et al. (2004), Franco and Filson (2006), Chatterji (2009)).[21] To assess the importance of this mechanism, in Table 7 we investigate the relationship between diversity and spin-out formation.

Because there is spin-out formation for only 4% of the MSA-class observations, our preferred specification is an OLS regression with MSA-class fixed effects, year effects, and technology class time trends.[22] The dependent variable is the logarithm of one plus the number of spin-outs.

The coefficient in Column (1) implies a 32% increase in the number of spin-outs when an MSA-class becomes diverse. In Column (2), we introduce additional controls for the number of labs in the MSA-class and their concentration. In this specification, the estimated diversity coefficient implies a 25% increase in spin-outs. Columns (3) and (4) confirm the result in the full sample that includes small MSA-classes. The estimates from these models show that diversity is associated with a 109% increase in the probability of spin-out formation. Column (5) shows that in the large sample, instrumenting diversity with lagged income taxes has essentially no impact on the estimated diversity coefficient.

Overall, the results in Table 7 provide direct evidence of a positive correlation between diversity and spin-out formation that is consistent with our theoretical framework. It is important to notice that these regressions exploit a restrictive measure of spin-out that relies only on patent data and requires inventors to patent both in large and new small labs. To provide additional support to the idea that spin-outs are the main mechanism through which diversity affects innovation, in the next sub-sections we present further indirect evidence consistent with our theory.

Firm Size Diversity and Non-Compete Agreements

Because in our theoretical framework the main channel through which diversity increases innovation is spin-out formation, an additional implication of our theory is that when spin-outs cannot be formed, the beneficial effect of diversity disappears and a single large lab maximizes innovation. This result suggests the correlation between innovation and

[21] Our data confirms this. For example, in 1995, spin-out patents receive 30% more citations than patents by other new entrants (average across all technology classes and MSAs)

[22] The conditional maximum likelihood estimation of the Poisson model drops MSA-classes in which the dependent variable is zero for the entire time period. This is equivalent to dropping about 84% of our data. With Poisson estimations in this smaller sample, we obtain a positive but statistically insignificant correlation between diversity and number of spin-outs.

size diversity is likely to be smaller in settings with substantial impediments to spin-out formation. To explore this idea empirically we interact *diverse* with the extent to which MSAs are located in states where non-compete agreements are strongly enforced. If high enforcement of non-compete laws prevents spin-out formation, then we expect the organization of the local R&D labor market to have no impact on innovation.

To construct a non-compete enforcement index, we follow Garmaise (2011) and construct a measure based on the twelve enforcement dimensions studied by Malsberger (2004). The index assigns one point to each dimension in which the jurisdiction law exceeds a given threshold; its value varies from zero to 12.

The original index constructed by Garmaise (2011) covers the period 1992-2005. We extend the time period from 1975 to 1992 by collecting information on changes in non-compete laws. The only change we identify for the period 1975-1992 occurs in Michigan in 1985.[23] Using this extended dataset, we construct a dummy variable *high enforce* that equals one when the index value is greater than or equal to six.[24]

We present regressions in which the diversity dummy is interacted with the high enforcement dummy in Table 8. As expected, we find negative and significant coefficients on the interaction terms, indicating that the effect of diversity is reduced by the presence of strong non-compete laws. The results are robust to introducing additional controls and to using the full sample that includes small MSAs.[25]

In interpreting these results, it is important to remember that, for simplicity, our theoretical model does not consider the competitive effect of spin-outs on large labs. In general, the effect of non-compete clauses will depend on the relationship between fit and competition. Low fit between an innovation and the existing lab activity does not imply an absence of competition between the spin-out and the large lab. Christensen (1997) provides a series of examples of companies not adopting new technologies because of opportunity costs (e.g., low initial profit margins or low attractiveness to the company's

[23]Prior to 1985, Michigan outlawed non-compete agreements, but in 1985 it passed legislation that enforced them. In the Garmaise (2011) data, the score for Michigan is five. We assign an index equal to zero before 1985 and an index of five after 1985.

[24]Following Stuart and Sorenson (2003) and Marx et al. (2010), we also generate a state-level indicator variable with a value of one if the state generally precludes, through statute or precedents, the enforcement of non-compete covenants. Also, with this measure, the results strongly support the predictions of the model.

[25]We cannot estimate the direct effect of $highenforce$ on innovation because the dummy is constant for all states over the entire sample period and thus collinear with the MSA-class fixed effects.

best customers) and eventually suffering from competing with these new technologies. If ideas that do not fit also have a competitive effect on large labs, then the innovation premium generated by diversity is lower when non-compete agreements are strongly enforced.[26]

Firm Size Diversity and Large Lab Focus

In our theoretical framework, spin-outs are formed to commercialize ideas that are not a good fit with the main research activity of a large lab. This implies that diversity has a greater impact on innovation the greater the number of misfit ideas. Empirically, this suggests an additional test to highlight the role of spin-outs as an important mechanism linking R&D market structure and innovation: diversity should have a differential impact on innovation depending on the likelihood that ideas fit in large lab research trajectories. We expect MSA diversity to have a greater impact on innovation when large labs follow narrow research trajectories and thus their ideas are more likely to be misfits. Similarly, when large labs follow broad research strategies, few spin-outs will be formed and the presence of small labs will be less beneficial to MSA innovation.[27]

---

[26]To see this, consider an innovation that in the case of a spin-out gives payoffs $\pi - k > 0$ to the new firm and $-z$ to the large lab. Without non-compete agreements, the large lab will have to pay $\pi - k$ to the inventor to prevent a spin-out and a spin-out will be formed as long as $\pi + a - (\pi - k) < -z$ or $a < -z - k$. If non-compete agreements can be enforced, then the large lab can stop the spin-out at zero cost and there will be no spin-out formation. In this case, the total number of commercialized innovations will be:

$$NI = \frac{S_L^2}{\Delta} H^S \left(1 - F(-\pi)\right) + \frac{N}{\Delta}, \tag{2}$$

which is maximized in the absence of small firms.

Implicitly, our analysis rules out transfers from scientists to large labs. In a Coasian framework where employees can pay their employers to leave the large lab, spin-outs may be formed even in the presence of non-compete agreements. The assumption of frictionless Coasian bargaining is hard to justify given the large evidence of transaction costs in technology transfer (Agrawal et al. (2012)).

[27]This prediction can be generated by our model assuming that a is distributed with cumulative distribution $F(a; b)$ where $b \in R$ generates a family of distributions ranked for first order stochastic dominance. For any b and b' if $b' \geq b$ we have that $F(a; b') \leq F(a; b)$. The parameter b captures the breadth of large lab research activity. If b is very low, then most ideas have low "fit" with the existing research activity of the lab. Let us compare two MSAs with the same number of scientists T. In the first "diverse" MSA, there is one large lab and $N' \geq \underline{N}$ small labs; in the second "non-diverse" MSA, there is one large lab but only $N < \underline{N}$ small labs. The difference in innovation between the diverse and the non-diverse MSA is equal to:

$$\frac{(T-N)^2}{\Delta} F(-\pi; b) + \frac{(N'-N)(N+N'+1-T)}{\Delta},$$

To construct a measure of narrowness for large lab research trajectories, for each large lab we compute the share of patents accounted for by the top four classes (C4 index) based on its patenting activity across three-digit technology classes (the 426 different USPTO n-classes) in a four-year window. For MSA-class-year cells with multiple large labs, we construct the mean value of the index for the labs in the cell. We generate an indicator variable, *large lab focus*, that equals one when the MSA-class-year C4 index is above the median relative to other MSAs within that class-year.

As an example of the variation in the data, consider chemical patenting activity in St. Louis, MO and San Diego, CA in 1994. Our data indicate the presence of only one large lab in both MSAs. For the St. Louis lab, the top four three-digit technology classes account for only 37% of its patenting, and the corresponding MSA-class large lab focus indicator is equal to zero. For the San Diego lab, the C4 index is approximately 85% and *large lab focus* is equal to one for this MSA-class.

In Table 9, we interact this variable with the diversity measure. Column (1) shows that the effect of diversity is stronger in MSA-classes where large labs follow narrow research trajectories. Columns (2) show that the correlation is robust to the inclusion of additional controls. Columns (3) and (4) show that results are similar in the larger sample that includes small MSAs.

Alternative Explanations and Robustness Checks

There are many reasons why co-existence of large and small firms may be associated with an increase in innovation. We first assess several such reasons in light of their consistency with the empirical correlations and provide additional results that help us rule out alternative theories that might be consistent with a subset, but not the whole set, of the correlations we report. We then describe a variety of additional extensions and robustness checks for interested readers in Appendix 2.

Product Market Competition

Aghion et al. (2005) develop a theoretical model predicting an inverted-U relationship between innovation and competition and present a series of empirical findings consistent

which is decreasing in b. This implies that diversity has greater impact on innovation when large labs follow narrow research trajectories (b is low).

with their theory. In light of their results, there is the concern that firm size diversity may simply capture intermediate levels of product market rivalry.

Notice that an important difference between Aghion et al. (2005) and our setting is that we conduct our analysis at the MSA-technology class level whereas their study is focused at the industry level. We expect most of the labs in our sample, especially the largest labs in diverse MSAs, to compete industry-wide at the national and international levels. We do not expect their product prices to be substantially affected by the local lab structure.

Moreover, the theoretical framework of Aghion et al. (2005) focuses on a duopoly and has no predictions on spin-out formation. The inverted U-shape theory is also difficult to reconcile with our findings that diversity has a greater impact when non-compete agreements are not strongly enforced and when large labs operate in few technology areas.

As a further robustness check, in Column (1) of Appendix Table A3, we introduce the square term for the Herfindahl concentration index. If the diversity dummy is only capturing intermediate values of concentration, introducing polynomial terms of the Herfindahl index should substantially reduce the correlation between diversity and innovation. Instead, we find that results are very similar when we introduce the square of the concentration index.[28]

Agglomeration Economies

Firm size diversity may be correlated with agglomeration economies that increase innovation productivity of labs in the MSA-class (Ellison et al. (2010)). In all our regressions, we control for the number of inventors working in the MSA-class, but this linear control can be inadequate if agglomeration economies arise only for very large MSAs.

Notice that if only agglomeration generates the diversity premium, then it is not clear why its impact is greater when non-compete agreements are not strongly enforced and when large labs operate in few technology areas. Nonetheless, to further address this concern, in Column (2) of Table A3 we introduce polynomial controls for the number of inventors in the MSA-class and the total number of inventors in the MSA. If the diversity dummy is only capturing agglomeration economies at MSA or MSA-class level, introducing these polynomial terms should substantially reduce the correlation between

[28]We obtain similar results using higher degree polynomials.

diversity and innovation. Instead, we find that results are very similar when we introduce the square terms.

In unreported regressions, we obtain similar results using higher degree polynomials. Results are also robust to controlling more flexibly for the number of MSA inventors by introducing a dummy for each decile of the distribution and there is essentially no change in the diversity coefficient in regressions that exploit the sample of large MSAs.

Large Lab Demand for Innovation and Licensees

Another competing explanation is that the role of large labs as consumers of small lab innovations drives the correlation between diversity and innovation. In other words, large labs generate a demand for technologies that induces entry by small firms and increases innovative output. Alternatively, the correlation may be driven by the role of small labs as licensees of large lab innovations that do not fit with large lab commercialization strategies.

While we expect these forces to have an impact on MSA innovation (Agrawal and Cockburn (2003)), alone these alternative theories cannot explain the entire set of correlations that we observe in the data. In particular, they are difficult to reconcile with the strong association between diversity and spin-out formation and with the finding that the effect of diversity is weaker when non-compete agreements are strongly enforced. Finally, we should expect large labs to boost innovation demand more strongly when they operate in a variety of technology areas, which is the opposite of what we find in our data.

Strategic Patenting

Ziedonis (2004), Lanjouw and Schankerman (2004), and Noel and Schankerman (2006) show that firms tend to expand their patent portfolios in response to potential hold-up problems generated by "thickets" in the market for technologies. If regional firm size diversity is associated with patent thickets, then the increase in innovation that we register may be due to the presence of overlapping and fragmented patent rights and not by the interplay of large and small firm externalities.

Previous studies, however, have documented the strategic patenting effect at the technology level and not at the regional level. Because our regressions include technology class time trends, they control for variation in "thickets" over time. Moreover, the

defensive patenting explanation would suggest an increase in the number of patents but a decline in the quality of the patents in the presence of regional diversity, which is the opposite of what we find in our data. Finally, strategic patenting is hard to reconcile with greater spin-out formation (in the same technology) and with the stronger impact of diversity of innovation when non-compete agreements are not enforced.

Innovation Affects Lab Structure

There is the concern that the correlation between diversity and innovation is due to changes in innovation outcomes rather than driven by changes in lab structure. In other words, "reverse causality" may take place if potential future innovation generates expansion of large labs and entry of small labs.

To address this concern, throughout our empirical analysis we use a five-year lag for the control variables. The IV regressions presented in Table 6 also address this concern by exploiting exogenous variation in MSA-class diversity. Finally, note that if an innovation shapes the lab configuration, then we must explain why its impact is greater when non-compete agreements are not strongly enforced and when large labs operate in few technology areas, neither of which are obvious.

Additional Robustness Checks

We conduct a number of additional empirical exercises to document the robustness of our main empirical results to alternative specifications and measurement strategies.

First, Appendix Table A2 provides evidence supporting the correlation between diversity and innovation in a series of regressions that distinguish between the direct effect of the presence of a large lab, the presence of numerous small labs, and their interaction. Columns (1) and (2) show that the presence of a large lab has no significant impact on innovation, whereas the presence of numerous small labs is associated with a greater number of cites per inventor. Columns (3) and (4) show that the two direct effects are not significant once we introduce the interaction between the two variables (the diversity dummy). These results confirm that the co-existence of at least one large lab and numerous small labs boosts innovation, not the presence of only one of these two factors.

There is concern that California MSAs, accounting for a large fraction of our switches in diversity, drive the results. In Column (3) of Appendix Table A3, we show

that we obtain similar results using the smaller sample of non-California MSAs.[29]

We also examine the technology specificity of the small firm externality. Specifically, we look at whether the diversity effect is driven by the presence of numerous small firms in the same MSA-class of a large lab compared to the presence of many small labs in the same MSA of a large lab (irrespective of class). Column (4) of Appendix Table A3 exploits a diversity measure that uses a small firm number cutoff at the MSA level rather than the MSA-class level. The correlation between this alternative diversity measure and innovation is small and not statistically significant, suggesting that externalities tend to be technology specific.[30]

We use patent data to construct the diversity measure, which may be associated with measurement error. To address this concern, we use an alternative diversity measure based on County Business Patterns Census data that reports detailed information on firm size across US MSAs. Because these data are available only for a sub-period of our panel, we focus on 1995 and construct a dummy (*diversecbp*) that equals one if the MSA-class has at least one establishment with more than 1,000 employees and a number of establishments with 5-49 employees above the 99th percentile. Column (5) of Table A3 shows that results are robust to using this alternative diversity measure.[31]

In Appendix Table A4, we show that results are similar if we employ different measures of diversity. We start by changing the definition of a large lab, moving its cutoff to the 98th and 99th percentile of the size distribution. We also alter the definition of a small lab, from exploiting all the labs below the 75th percentile to using labs in the 50th-75th range and the 50th-97th range. Results are similar if we reduce the threshold level of small labs in the diversity measure by 20% and 30% (from 139 to 111 and 97, respectively).

In unreported regressions, we also alter the definition of the large MSA sample. In the previous tables, we use the entire sample and MSA-class years where the number of inventors is above 471 (the minimum required to have diversity). Results are robust to

[29]To conduct this exercise, we rescale the diversity measure using the distributions of large lab size and of small firms in this new sample.

[30]We also examine whether MSA-classes with numerous small firms benefit from the existence of large labs operating in the same technology area but located in different MSAs in the same state. We find that this is not the case: co-existence of large and small labs is associated with an innovation premium only if large and small labs are located in the same MSA-class.

[31]We obtain similar results using spin-outs as the dependent variable. There is no change in the diversity coefficient if we control for the Census measure of the number of employees in the MSA-class.

considering MSA-classes that have more than 471 inventors for the entire sample period or to having an MSA-class entering the sample when it passes the 471 threshold. Finally, results are robust to dropping technology class 6 (Miscellaneous).

6 Conclusion

Our results suggest that the way in which R&D labor is organized in a region is associated with its dynamism and growth, at least insofar as this is captured by rates of patenting. Our findings point to a potentially important role for market structure in driving the performance of local innovation economies. Like Aghion et al. (2005) but on a different dimension (R&D labor market rather than product market), we find that extreme structures are not optimal for innovation. Policies focused exclusively either on attracting "anchor tenants" or on cultivating small entrepreneurial ventures may therefore be missing an important opportunity.[32]

Following an approach that is common in the empirical literature on the determination of industries' innovative activity (Aghion et al. (2005), Bloom et al. (2010), Cohen (2010)), we examine the interplay between firm size diversity and innovation by performing comparative statics in the regional organization of R&D labor. Because innovation and regional lab structure are mutually endogenous, our empirical analysis exploits fixed effects, lagged dependent variables, and instrumental variables to identify the impact of diversity on innovation. We leave for future research an analysis of the impact of innovation on R&D labor structure, wages, and their evolution over time.

Our results also point to additional opportunities and challenges for further research. While our characterization of market structure is richer than one-dimensional measures such as the Herfindahl index or concentration ratios, it still represents only a first step. We note that even on the somewhat easier turf of the product market, decades of empirical research have yet to establish robust relationships between traditional measures of market structure and outcomes such as price-cost margins. Furthermore, while we distinguish between incumbents' and entrants' contributions to overall innovation activity, this represents only initial progress towards understanding the impact of local

[32] An additional implication of our results is that policies focused on facilitating spin-out formation (e.g., amending non-compete enforcement laws) may also have a beneficial effect on regional innovation. It remains for future research to study how different policies may interact with the organization of R&D labor organization and affect regional innovation.

innovation market structure on the intensive versus extensive margin of innovation by different types of firms.

Finally, our empirical analysis is sparing in its use of explanatory variables and also relies heavily on only one noisy measure of innovative output, patents. We would prefer to capture other factors that potentially directly affect innovation performance, such as Glaeser-type "amenities," rather than through fixed effects. Linking the types of patent statistics used here to other sources, such as demographics or data on the production economy, may help with external validation as well as provide the basis for a richer investigation of local innovation markets.

References

Agarwal, R., R. Echambadi, A. M. Franco, and M. Sarkar (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *The Academy of Management Journal 47*(4), pp. 501–522.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005, May). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics 120*(2), 701–728.

Agrawal, A. and I. Cockburn (2003, November). The anchor tenant hypothesis: exploring the role of large, local, r&d-intensive firms in regional innovation systems. *International Journal of Industrial Organization 21*(9), 1227–1253.

Agrawal, A., I. Cockburn, and L. Zhang (2012). Deals not done: Sources of failure in the market for ideas. University of Toronto Working Paper.

Agrawal, A. and A. Goldfarb (2008). Restructuring research: Communication costs and the democratization of university innovation. *American Economic Review 98*(4), 1578–90.

Arora, A. and A. Gambardella (1994). The changing technology of technological change: General and abstract knowledge and the division of innovative labour. *Research Policy 23*(5), 523–532.

Bloom, N., C. Propper, S. Sailer, and J. V. Reenen (2010). The impact of competition on management quality: Evidence from public hospitals. NBER Working Paper # 16032.

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

Bresnahan, T., S. Greenstein, and R. Henderson (2011). Schumpeterian competition and diseconomies of scope; illustrations from the histories of microsoft and ibm. *HBS working paper 11-07*, 362–78.

Bruce, D. (2000). Effects of the united states tax system on transitions into self-employment. *Labour Economics 7*, 545–574.

Cassiman, B. and M. Ueda (2006). Optimal project rejection and new firm start-ups. *Management Science 52*, 262–75.

Chatterji, A. K. (2009). Spawned with a silver spoon? entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal 30*(2), 185–206.

Chinitz, B. (1961). Contrasts in agglomeration: New york and pittsburgh. *The American Economic Review 51*(2), pp. 279–289.

Christensen, C. (1997). *Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Press Books.

Cockburn, I. M. and R. M. Henderson (2001, November). Scale and scope in drug development: Unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics 20*(6), 1033–1057.

Cohen, W. M. (2010). Fifty years of empirical studies of innovative activity and performance. In B. H. Hall and N. Rosenberg (Eds.), *Handbook of The Economics of Innovation, Vol. 1*, pp. 129 – 213. North-Holland.

Cohen, W. M. and S. Klepper (1996, July). A reprise of size and r&d. *Economic Journal 106*(437), 925–51.

Cullen, J. B. and R. H. Gordon (2007). Taxes and entrepreneurial risk-taking: Theory and evidence for the u.s. *Journal of Public Economics 91*, 1479–1505.

Ellison, G., E. Glaeser, and W. Kerr (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review 100*(3), 1195–1213.

Feenberg, D. and E. Coutts (1993). An introduction to the taxsim model. *Journal of Policy Analysis and Management 12*(1), 189–194.

Feldman, M. P. and D. B. Audretsch (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review 43*(2), 409 – 429.

Franco, A. and D. Filson (2006). Spin-outs: knowledge diffusion through employee mobility. *The RAND Journal of Economics 37*(4), 841–860.

Galasso, A., M. Schankerman, and C. Serrano (2011). Trading and enforcing patent rights. NBER working paper #17367.

Gans, J. and S. Stern (2000). Incumbency and r&d incentives: Licensing the gale of creative destruction. *Journal of Economics and Management Strategy 9*, 485–511.

Garmaise, M. J. (2011). Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *Journal of Law, Economics, and Organization 27*(2), 376–425.

Gentry, W. and G. Hubbard (2004). Success taxes, entrepreneurial entry and innovation. NBER working paper #10551.

Geroski, P. A. (1990, July). Innovation, technological opportunity, and market structure. *Oxford Economic Papers 42*(3), 586–602.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992, December). Growth in cities. *Journal of Political Economy 100*(6), 1126–52.

Glaeser, E. L. and W. R. Kerr (2009, 09). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics & Management Strategy 18*(3), 623–663.

Grossman, G. M. and E. Helpman (1991, May). Quality ladders and product cycles. *The Quarterly Journal of Economics 106*(2), 557–86.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001, October). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Working Paper 8498, National Bureau of Economic Research.

Hausman, J., B. Hall, and Z. Griliches (1984). Econometric Models for Count Data with an Application to the Patents-R&D Relationship. *Econometrica 52*(4), 909–938.

Helsley, R. W. and W. C. Strange (2002). Innovation and input sharing. *Journal of Urban Economics 51*, 25–45.

Henderson, R. and I. Cockburn (1996, Spring). Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *RAND Journal of Economics 27*(1), 32–59.

Jacobs, J. (1969). *The Economy of Cities.* Random House: New York.

Jones, B. (2008). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies 76*(1), 283–317.

Klepper, S. and P. Thompson (2010). Disagreement and Intra-industry spinoffs. *International Journal of Industrial Organization 28*, 526–538.

Klette, T. J. (1996, Autumn). R&d, scope economies, and plant performance. *RAND Journal of Economics 27*(3), 502–522.

Lanjouw, J. O. and M. Schankerman (2004, April). Protecting intellectual property rights: Are small firms handicapped? *Journal of Law and Economics 47*(1), 45–74.

Long, J. (1982). The income tax and self-employment. *National Tax Journal 35*, 31–42.

Malsberger, B. M. (2004). *Covenants Not to Compete : A State-by-State Survey*. Washington, DC: BNA Books.

Marx, M., J. Singh, and L. Fleming (2010). Does non-compete enforcement create a brain drain? HBS working paper.

Noel, M. and M. Schankerman (2006). Strategic patenting and software innovation. CEPR Discussion Paper 5701.

Pakes, A. and Z. Griliches (1980). Patents and R&D at the firm level: A first report. *Economics Letters 5*(4), 377–381.

Schumpeter, J. A. (1942). *Capitalism, Socialism, and Democracy*. New York, Harper.

Stuart, T. E. and O. Sorenson (2003). Liquidity events and the geographic distribution of entrepreneurial activity. *Administrative Science Quarterly 48*(2), pp. 175–201.

Vernon, R. (1960). *Metropolis 1985*. Harvard University Press Cambridge, MA.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics 90*(1), 77–97.

Ziedonis, R. H. (2004). Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Science 50*(6), pp. 804–820.

Figure 1: Variation in Regional Innovation

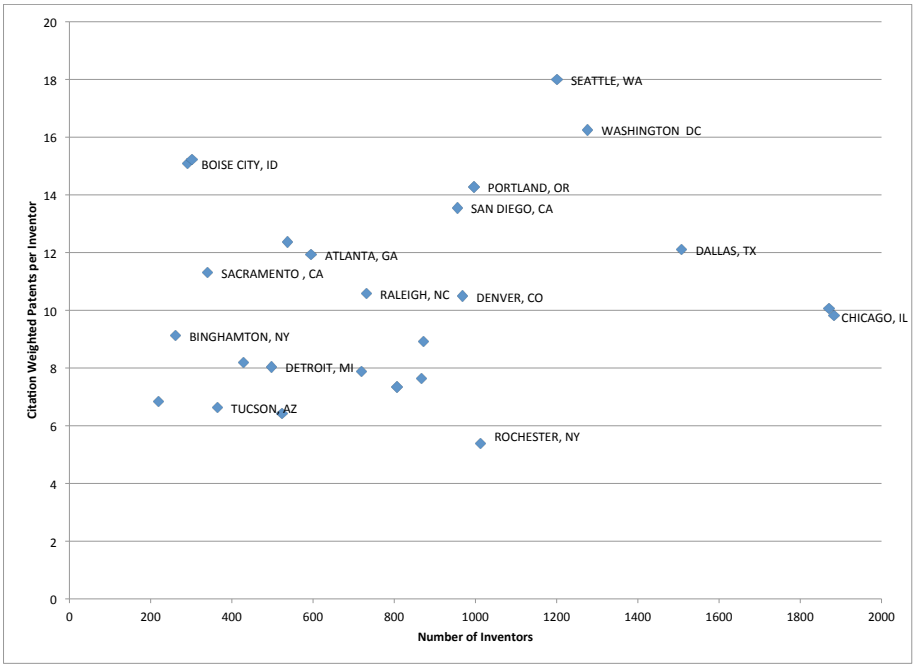

Figure 2: Commercialization of Large Lab Innovations

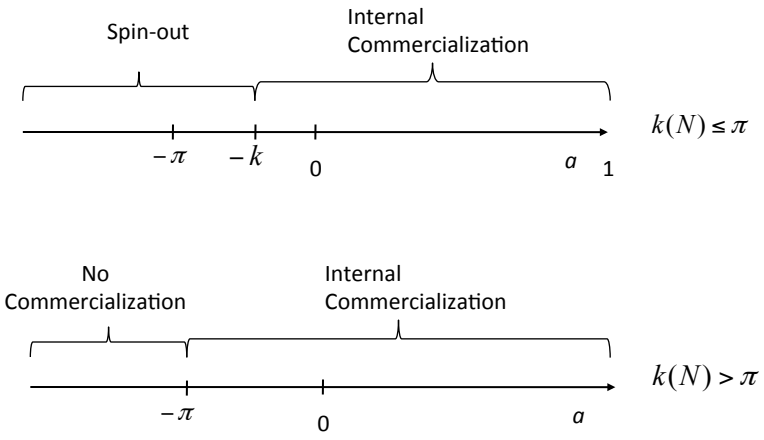

Table 1: Summary Statistics; N = 41,808

Variables	mean	median	Std. Dev.	min	max
Weighted Patents$_{jkt+5}$	465.72	33	2556.94	0	161861
Patents$_{jkt+5}$	27.73	3	118.42	0	6436
Diverse	0.01	0	0.08	0	1
LargeLab	0.11	0	0.32	0	1
# Small Labs	6.43	1	23.37	0	995
# Inventors	91.16	5	411.66	0	23689
# LargeLab Inventors	33.62	0	199.81	0	11725
# Labs	8.08	1	29.76	0	1318
Herf	0.27	0	0.38	0	1
Spin-out	0.08	0	0.26	0	1
# Spin-outs	0.18	0	1.80	0	158
High Enforce	0.27	0	0.44	0	1
No Enforce	0.21	0	0.41	0	1
Large Lab Focus	0.06	0	0.23	0	1

Table 2: Summary Statistics; N = 1,870

Variables	mean	median	Std. Dev.	min	max
Weighted Patents$_{t+5}$	6680.68	4166.0	10093.74	418.00	161861
Patents$_{t+5}$	381.71	268.0	413.65	72.00	6436
Diverse	0.16	0.0	0.36	0.00	1
LargeLab	0.96	1.0	0.20	0.00	1
# Small Labs	83.03	63.0	73.10	2.00	995
# Inventors	1347.22	905.0	1425.11	472.00	23689
# LargeLab Inventors	574.62	366.5	747.89	0.00	11725
# Labs	105.08	80.0	94.01	2.00	1318
Herf	0.15	0.1	0.16	0.01	1
Spin-out	0.57	1.0	0.50	0.00	1
# Spin-outs	2.35	1.0	5.62	0.00	111
High Enforce	0.24	0.0	0.43	0.00	1
No Enforce	0.30	0.0	0.46	0.00	1
Large Lab Focus	0.45	0.0	0.50	0.00	1

Table 3: Variation in Diversity (Year of the Switch)

MSA	Chemical	Computers	Drugs	Electronics	Mechanical	Other
Boston	1993	1995	1994	1992	1996	1996
Chicago	1996	2000			1979	1978
Dallas		1999				
Detroit					1994	
LA	1998	1995	1995	1993	1978	1980
NYC	1976	1995	1989	1987	1978	1977
Philadelphia	1996		1998			
San Diego		2000	1998			
San Francisco	1990	1989	1992	1987	1995	1994
Seattle		2000				
DC		1998				

Table 4: Diversity and Innovation (Cross-Section)

	(1)	(2)	(3)	(4)
Sample:	Full, 1995	Full, 1995	Full, 1995	Full, 1995
Dependent Variable:	Weighted Patents$_{jkt=2000}$	Weighted Patents$_{jkt=2000}$	Weighted Patents$_{jkt=2000}$	Patents$_{jkt=2000}$
diverse$_{jkt=1995}$	3.762**	0.406*	0.382*	0.296**
	(0.333)	(0.185)	(0.165)	(0.106)
logInventors$_{jkt=1995}$		0.937**	0.891**	0.861**
		(0.050)	(0.027)	(0.017)
Constant	6.005**	1.841**	1.348**	−0.328**
	(0.140)	(0.284)	(0.140)	(0.109)
Class FE			✓	✓
Observations	1608	1608	1608	1608
log likelihood	−1260012.43	−303371.15	−168774.27	−12157.30

Notes: Observations are at the MSA$_j$-class$_k$ level. All specifications are Poisson regressions estimated by maximum likelihood. *Weighted patents* is the forward citation weighted sum of distinct patents with primary technology classification k and application year t + 5 where at least one inventor is located in MSA j. The main independent variable, *diverse*, equals 1 if MSA j in class k in 1995 has at least one large lab and more than 139 active small labs, and 0 otherwise. *Loginventors* is the log of the number of distinct active inventors in MSA j, class k in 1995.
Robust standard errors clustered at the MSA-class level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 5: Diversity and Innovation

Sample:	(1) Full	(2) At-Risk	(3) At-Risk	(4) Full	(5) Full	(6) Full
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Patents$_{jkt+5}$
diverse$_{jkt}$	0.309**	0.161**	0.137**	0.146*	0.127*	0.085*
	(0.116)	(0.053)	(0.050)	(0.063)	(0.058)	(0.043)
logInventors$_{jkt}$	0.911**	0.644**	0.518**	0.293**	0.200**	0.334**
	(0.014)	(0.077)	(0.097)	(0.034)	(0.054)	(0.040)
logNumLabs$_{jkt}$			0.187		0.205**	0.061
			(0.138)		(0.071)	(0.052)
herf$_{jkt}$			−0.763		−0.074	−0.108**
			(0.563)		(0.052)	(0.035)
Year X Class FE	✓	✓	✓	✓	✓	✓
MSA X Class FE		✓	✓	✓	✓	✓
Observations	41808	1864	1864	41262	41262	41444
Num. Groups		140.00	140.00	1587.00	1587.00	1594.00
log likelihood	−3781751.18	−143844.93	−140280.30	−1345383.49	−1332410.79	−120051.25

Notes: Observations are at the MSA$_j$-class$_k$-year$_t$ level. All specifications are estimated with Poisson with Fixed Effects Quasi maximum likelihood (QML).
Weighted patents is the forward citation weighted sum of distinct patents with primary technology classification k and application year t + 5 where at least one inventor is located in MSA j. The main independent variable, *diverse*, equals 1 if MSA j in class k in year t has at least one large lab and more than 139 active small labs, and 0 otherwise. *Loginventors* is the log of the number of distinct active inventors in MSA j, class k in year t, *logNumLabs* is the log of the number of distinct active labs in MSA j, class k in year t, and *herf* is the Herfindahl concentration index for labs in MSA j, class k in year t. Column 3 and 4 only includes MSA-class-years that have enough inventors to be at-risk of displaying a positive diversity measure.
Robust standard errors clustered at the MSA-class level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 6: Innovation and Lagged Income Tax IV

	(1)	(2)	(3)	(4)
Sample:	At-Risk	At-Risk	At-Risk	At-Risk
Estimation:	OLS	First Stage	2SLS	2SLS
Dependent Variable:	ln Weighted Patents$_{jkt+5}$	diverse$_{jkt}$	ln Weighted Patents$_{jkt+5}$	ln Weighted Patents$_{jkt+5}$
diverse$_{jkt}$	0.180**		0.683**	0.836**
	(0.048)		(0.236)	(0.215)
logInventors$_{jkt}$	0.525**	0.229**	0.300**	0.315**
	(0.076)	(0.035)	(0.069)	(0.070)
logIncomeTax$_{jt}$		0.194	0.825*	0.760
		(0.403)	(0.419)	(0.593)
logIncomeTax$_{jt-3}$		1.402**		
		(0.340)		
logGainsTax$_{jt}$				0.032
				(0.031)
Year X Class FE	✓	✓	✓	✓
MSA X Class FE	✓	✓	✓	✓
Observations	1873	1702	1693	1517
Num. Groups	149.00	149.00	140.00	121.00
First Stage F-statistic		17.53		
R^2	0.94	0.41	0.92	0.91

Notes: Column 1 replicates the specification estimated in Table 3, Column 4 with OLS. This specification serves as a baseline for results when estimated with linear least squares. Column 2 presents the first stage regression with *diverse* as the dependent variable. *logIncomeTax* is the log of personal income tax in year t in MSA j's state. Column 3 presents 2SLS results with diverse instrumented by $logIncomeTax_{t-3}$. Column 4 includes the additional control $logGainsTax_{t-3}$, the log of capital gains tax in year $t-3$ in MSA j's state.
Robust standard errors clustered at the MSA-class level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 7: Diversity and Spin-Out Formation

	(1)	(2)	(3)	(4)	(5)
Sample:	At-Risk	At-Risk	Full	Full	Full
Estimation:	OLS	OLS	OLS	OLS	2SLS
Dependent Variable:	ln Spin-outs$_{jkt}$	ln Spin-outs$_{jkt}$	ln Spin-outs$_{jkt}$	ln Spin-outs$_{jkt}$	ln Spin-outs$_{jkt}$
diverse$_{jkt}$	0.324**	0.246**	1.097**	1.096**	1.066**
	(0.071)	(0.070)	(0.110)	(0.110)	(0.321)
logInventorsLarge$_{jkt}$	0.071**	0.007	0.063**	0.061**	0.059**
	(0.025)	(0.019)	(0.004)	(0.004)	(0.004)
logInventors$_{jkt}$		0.638**		0.007**	0.010**
		(0.111)		(0.002)	(0.001)
logIncomeTax$_{jt}$					−0.041
					(0.065)
Year X Class FE	✓	✓	✓	✓	
MSA X Class FE	✓	✓	✓	✓	
Observations	1873	1873	41808	41808	30552
Num. Groups	149.00	149.00	1608.00	1608.00	1608.00
First Stage F-statistic					91.24
R^2	0.49	0.51	0.23	0.23	0.19

Notes: The dependent variable, *ln Spin-outs*$_{jkt}$ is the logarithm of one plus the count of spin-outs in MSA j, class k in year t. The variable *diverse*$_{jkt}$ in Column 5 is instrumented with *logIncomeTax*$_{jt-3}$ and *logIncomeTax*$_{jt-5}$. Because of the five-year lag in these instruments, the panel in Column 5 includes five fewer years.
Robust standard errors clustered at the MSA-class level in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 8: Innovation and Non-compete Clauses

	(1)	(2)	(3)	(4)
Sample:	At-Risk	At-Risk	Full	Full
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$
diverse$_{jkt}$	0.189**	0.168**	0.175*	0.157*
	(0.064)	(0.061)	(0.072)	(0.066)
diverse$_{jkt}$ X high enforce$_j$	−0.152$^+$	−0.174*	−0.194*	−0.206**
	(0.091)	(0.082)	(0.084)	(0.079)
logInventors$_{jkt}$	0.630**	0.475**	0.291**	0.197**
	(0.078)	(0.101)	(0.034)	(0.054)
logNumLabs$_{jkt}$		0.227		0.209**
		(0.139)		(0.070)
herf$_{jkt}$		−0.730		−0.073
		(0.546)		(0.052)
Year X Class FE	✓	✓	✓	✓
MSA X Class FE	✓	✓	✓	✓
Observations	1864	1864	41262	41262
Num. Groups	140.00	140.00	1587.00	1587.00
log likelihood	−142174.66	−138097.89	−1342338.58	−1328971.60

Notes: All specifications are estimated by Poisson with Fixed Effects QML. The variable *high enforce* is equal to 1 if MSA j is in a high-enforcement state in year t. A state is high-enforcement if it has a value of 5 or greater on the 13-point scale (0-12) based on Garmaise (2011). This variable is time-invariant during our sample period, and so its effect is identified through the interaction with the time-varying variable diverse.
Robust standard errors clustered at the MSA-class level in parentheses.
$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 9: Innovation, Large Lab Focus and Diversity

Sample:	(1) At-Risk	(2) At-Risk	(3) Full	(4) Full
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$
diverse$_{jkt}$	0.118*	0.105^{+}	0.077	0.071
	(0.060)	(0.055)	(0.062)	(0.059)
large lab focus$_{jkt}$	−0.028	−0.020	−0.007	0.005
	(0.030)	(0.029)	(0.031)	(0.029)
diverse$_{jkt}$ X large lab focus$_{jkt}$	0.096*	0.077*	0.146**	0.121*
	(0.041)	(0.039)	(0.055)	(0.050)
logInventors$_{jkt}$	0.613**	0.507**	0.284**	0.190**
	(0.070)	(0.095)	(0.032)	(0.051)
largelab$_{jkt}$	0.052	0.047	0.019	0.029
	(0.069)	(0.066)	(0.037)	(0.035)
logNumLabs$_{jkt}$		0.164		0.203**
		(0.133)		(0.066)
herf$_{jkt}$		−0.733		−0.070
		(0.579)		(0.051)
Year X Class FE	✓	✓	✓	✓
MSA X Class FE	✓	✓	✓	✓
Observations	1864	1864	41262	41262
Num. Groups	140.00	140.00	1587.00	1587.00
log likelihood	−142142.56	−139157.24	−1339608.32	−1327459.97

Notes: All specifications are estimated by Poisson with Fixed Effects QML. The variable *large lab focus* is equal to 1 if the C4 concentration ratio of the patenting activity across three-digit technology classes of the large labs in MSA j, class k in year t is above the median value. This measure captures MSAs whose large labs patent in more focused (narrow) technology areas.
Robust standard errors clustered at the MSA-class in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Appendix 1: Theoretical Results

Proof of Proposition 1

In a small lab, it is efficient to commercialize an innovation if $\pi \geq 0$. To avoid spin-out formation, large labs need to offer the scientist $\pi - k(N)$. This implies that spin-outs are formed if $\pi + a - (\pi - k(N)) \leq 0$ or $a \leq -k(N)$.

Proof of Proposition 2

We denote with NS the expected number of spin-outs in the MSA.

$$\begin{aligned} NS &= \sum_{j=1}^{J} \frac{S_j^2}{\Delta} \Pr\textit{(spin-out)} \\ &= \begin{cases} \dfrac{S_L^2}{\Delta} H^S F\left(-k(N)\right) & \text{if} \quad N \geq \underline{N} \\ \\ 0 & \text{if} \quad N < \underline{N} \end{cases} . \end{aligned} \tag{3}$$

where $H^S = \sum_{j=1}^{J} \left(S_j/S_L\right)^2$ is the Herfindahl concentration index of large labs.

Differentiation of (3) shows that NS is maximized when $H^S = 1$, which implies that at most one large lab will be present. Because there are no spin-outs if $N < \underline{N}$, the number of small labs that maximizes spin-out formation is the largest between $\underline{N}$ and N^* where we define N^* as

$$N^* = \arg\max_N \varphi(N)$$

where $\varphi(N) \equiv \dfrac{(T-N)^2}{\Delta} F\left(-k(N)\right)$.

We now consider the total innovation activity in the MSA. The total number of

commercialized innovations is:

$$\begin{aligned} NI &= \sum_{j=1}^{J} \frac{S_j^2}{\Delta} \Pr(\textit{Comm. from Large}) + \frac{N}{\Delta} \\ &= \begin{cases} \dfrac{S_L^2}{\Delta} H^S + \dfrac{N}{\Delta} & \text{if} \quad N \geq \underline{N} \\ \dfrac{S_L^2}{\Delta} H^S \left(1 - F(-\pi)\right) + \dfrac{N}{\Delta} & \text{if} \quad N < \underline{N} \end{cases} . \end{aligned} \tag{4}$$

NI increases in H^S so it is maximized when $H^S = 1$. Consider now the case in which $N < \underline{N}$. Evaluated at $H^S = 1$, NI has the following functional form:

$$g(N) = \frac{(T-N)^2}{\Delta} \left(1 - F(-\pi)\right) + \frac{N}{\Delta}$$

that is decreasing in N for T large enough. When $N \geq \underline{N}$ and $H^S = 1$, NI can be written as:

$$z(N) = \frac{(T-N)^2}{\Delta} + \frac{N}{\Delta}$$

that is also decreasing in N. It is easy to see that $z(N) > g(N)$, so innovation is maximized with diversity if $z(\underline{N}) > g(0)$, which we can rewrite as:

$$(T - \underline{N})^2 + \underline{N} > T^2(1 - F(-\pi))$$

or

$$T\left[TF(-\pi) - 2\underline{N}\right] > -\underline{N}^2 - \underline{N}$$

that is satisfied for T large enough.

Comparison of Diverse and Non-Diverse MSAs

Take an arbitrary "diverse" MSA with $N \geq \underline{N}$ small labs and one large lab of size $(T - N)$. We first show that this configuration generates more commercialized ideas than any non-diverse MSA with $N < \underline{N}$. Notice that when $N < \underline{N}$, we maximize the amount of commercialized ideas by allocating all scientists to a single large lab: $T^2(1 - F(-\pi))/\Delta$. This amount of commercialized ideas is lower than the one in the diverse MSA if:

$$(T - N)^2 + N > T^2(1 - F(-\pi))$$

that is satisfied for T large enough. Consider now a non-diverse MSA with $N \geq \underline{N}$ and multiple large labs, each of a size not exceeding $T - N$. Among all these MSAs, we maximize the amount of commercialized innovation by allocating $(T - N - 1)$ to one large lab, $(N - \underline{N})$ scientists to a second large lab, and $\underline{N}$ scientists to small labs. The amount of commercialized ideas of this MSA is lower than the one in the diverse MSA if:

$$\begin{aligned} (T-N)^2 + N &> (T-N-1)^2 + (N-\underline{N})^2 + \underline{N} \\ 2(T-N) &> (N-\underline{N})^2 - (N-\underline{N}) + 1 \end{aligned}$$

that is satisfied when T is large enough. This shows that the diverse MSA dominates all the non-diverse MSAs with $N \geq \underline{N}$ and multiple large labs, each of a size not exceeding $T - N$.

We now consider spin-out formation. Because no spin-out takes place if $N < \underline{N}$, the diverse MSA generates more spin-outs than all the non-diverse MSA with $N < \underline{N}$. Consider now a non-diverse MSA with $\widetilde{N} \in [N, \underline{N}]$ small labs and multiple large labs, each of a size not exceeding $T - N$. The MSA generates the maximum amount of spin-outs when we allocate $(T - N - 1)$ scientists to one large lab and $\left(N - \widetilde{N}\right)$ scientists to a second large lab. This configuration will generate less spin-outs than a diverse MSA if

$$(T-N)^2 F(-k(N)) \geq \left[(T-N-1)^2 + (N-\widetilde{N})^2\right] F(-k(\widetilde{N})$$

that is satisfied for T large enough because $F(-k(N)) \geq F(-k(\widetilde{N})$. This shows that the diverse MSA dominates all the non-diverse MSAs with $\widetilde{N} \in [N, \underline{N}]$ small labs.

Relaxing the Definition of Diversity

Take an arbitrary "diverse" MSA with $N \geq \underline{N}$ small labs and multiple large labs of size S_j with $j = 1, ..., J$. We first show that this configuration generates more commercialized ideas than any non-diverse MSA with $N < \underline{N}$. Notice that when $N < \underline{N}$, we obtain the maximum amount of commercialized ideas by allocating all scientists to a single large lab: $T^2(1 - F(-\pi))/\Delta$. The amount of commercialized ideas is lower than the one in the

diverse MSA if:

$$\sum_j S_j^2 + N > T^2(1 - F(-\pi)). \tag{5}$$

Call k the largest lab and notice that ${S_k}^2 + N < \sum_j S_j^2 + N$. Then a sufficient condition to have (5) satisfied is:

$${S_k}^2 + N > T^2(1 - F(-\pi))$$

that is satisfied if $S_k > \underline{S} \equiv \sqrt{T^2(1 - F(-\pi)) - N}$.

Appendix 2: Additional Tables

Table A.1: Income Tax Rates

Period	Mean	Std. Dev	Min.	Max.
1977-1979	52.80	2.26	50	59.9
1980-1984	56.68	1.95	50	56.9
1985-1989	42.02	9.67	28	56.25
1990-1994	38.27	5.46	28	48.15
1995-1999	44.03	1.84	40.79	46.89
2000-2005	40.99	2.63	36.05	46.28

Table A.2: Innovation and Diversity - Robustness with Levels

	(1)	(2)	(3)	(4)
Sample:	Full	Full	Full	Full
Estimation:	QML	QML	QML	QML
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$
largelab$_{jkt}$	−0.002		0.012	0.031
	(0.039)		(0.035)	(0.032)
Smallthresh$_{jkt}$		0.146*	−0.080	−0.064
		(0.064)	(0.068)	(0.067)
diverse$_{jkt}$			0.227**	0.193*
			(0.083)	(0.077)
logInventors$_{jkt}$	0.288**	0.292**	0.290**	0.191**
	(0.034)	(0.034)	(0.033)	(0.051)
logNumLabs$_{jkt}$				0.211**
				(0.070)
herf$_{jkt}$				−0.063
				(0.051)
Year X Class FE	✓	✓	✓	✓
MSA X Class FE	✓	✓	✓	✓
Observations	41262	41262	41262	41262
Num. Groups	1587	1587	1587	1587
log likelihood	−1356561.51	−1346174.30	−1346108.58	−1333246.93

Notes: Robust Standard errors clustered at MSA-class in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A.3: Additional Robustness Checks I

Sample:	(1) Full	(2) Full	(3) No California‡	(4) Full	(5) Full, 1995
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt=2000}$
diverse$_{jkt}$	0.123*	0.119*	0.095**		
	(0.057)	(0.052)	(0.035)		
diversemsa$_{jt}$				0.031	
				(0.042)	
diversecbp$_{jkt=1995}$					0.279**
					(0.099)
logInventors$_{jkt}$	0.225**	0.192**	0.250**	0.286**	0.916**
	(0.054)	(0.058)	(0.039)	(0.034)	(0.043)
logNumLabs$_{jkt}$	0.177*				
	(0.071)				
herf$_{jkt}$	−0.496**				
	(0.18)				
herf$^2_{jkt}$	0.407*				
	(0.162)				
logInventors$^2_{jkt}$		−0.010+			
		(0.006)			
logMSAInventors$_{jt}$		0.182**			
		(0.052)			
logMSAInventors$^2_{jt}$		0.018**			
		(0.006)			
Year X Class FE	✓	✓	✓	✓	
MSA X Class FE	✓	✓	✓	✓	
Class FE					✓
Observations	41262	41262	38792	41262	1608
Num. Groups	1587.00	1587.00	1492.00	1587.00	
log likelihood	−1331011.70	−1290636.07	−1193672.34	−1356126.58	−172699.85

Notes: *diversecbp$_t$* corresponds to a diversity measure constructed with data from the Census County Business Patterns dataset (1995). *logMSAInventors$_t$* corresponds to the log of inventors in MSA j across all classes k.
‡ We construct *diverse* in this specification using the rescaled distribution of firms outside California.
Robust standard errors clustered at the MSA-class in parentheses.
$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table A.4: Additional Robustness Checks II

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Full	Full	Full	Full	Full	Full
Diverse Definition:	LargeLabs $> 98^{th}$ p	LargeLabs $> 99^{th}$ p	Small Labs 50^{th}-75^{th} p	Small Labs 50^{th}-97^{th} p	# Small Labs > 111	# Small Labs > 97
Dependent Variable:	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$	Weighted Patents$_{jkt+5}$
diverse$_{jkt}$	0.147^{*}	0.134^{*}	0.140^{*}	0.135^{*}	0.116^{+}	0.110^{+}
	(0.064)	(0.068)	(0.058)	(0.061)	(0.062)	(0.058)
logInventors$_{jkt}$	0.292^{**}	0.292^{**}	0.291^{**}	0.291^{**}	0.291^{**}	0.290^{**}
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Year X Class FE	✓	✓	✓	✓	✓	✓
MSA X Class FE	✓	✓	✓	✓	✓	✓
Observations	41262	41262	41262	41262	41262	41262
Num. Groups	1587.00	1587.00	1587.00	1587.00	1587.00	1587.00
log likelihood	−1346082.01	−1348483.79	−1346467.99	−1347062.73	−1348758.11	−1349537.11

Notes: Robust Standard errors clustered at MSA-class in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$