Research Policy xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

# Triangulating regional economies: Realizing the promise of digital data

Maryann Feldman*, Nichola Lowe

*University of North Carolina, Chapel Hill, USA*

ARTICLE INFO

*Article history:*
Received 23 January 2015
Accepted 24 January 2015
Available online xxx

*Keywords:*
Geography of innovation
Entrepreneurship
Regional dynamics
Organizational change

ABSTRACT

Innovative data sources offer new ways of studying spatial and temporal industrial and regional development. Our approach is to study the development of an entrepreneurial regional economy through a comprehensive analysis of its constituent firms and institutions over time. Our study region is defined by the location of large multinationals recruited to North Carolina's Research Triangle Park and the adjacent area. We have built a database of 4200 technology-intensive entrepreneurial firms that draws on over 30 distinct data sources and includes details on company founders, annual firm employment and engagement with the entrepreneurial ecosystem. By outlining our approach in this paper, our primary objective is to create a transferable framework for analyzing regional dynamics in other locations.

Published by Elsevier B.V.

## 1. Introduction

Digital data provides a means to understand the functioning of innovative industries and regional economies. Aggregate industrial data produced by a variety of government agencies and dependent on rigid and inflexible geographic units, while useful for certain analyses, obscures many research questions relevant to studying the evolution of regional economies. Consideration of the role of firms—as dynamic actors with evolving capabilities and temporally delineated strategic investments in a region—is frequently absent because of the restrictions to sharing and collecting micro-level data. Moreover, comparative analysis typically assumes static or time-invariant institutional features and therefore is unable to capture the processes by which organizations and institutions within a region change and adapt over time. Thus, we are left to assume that regional institutional features are deterministic and static rather than socially constructed. Moreover, outcomes we observe are the result of a complex set of actors and specific interactions that occur in a temporal and varying regional setting. An unfortunate disconnect exists between the theoretical definition of region as integrated contiguous space and the political and Census geography for which data are readily available. Insufficient micro-level data has inhibited understanding of the underlying dynamic processes within regions that lead to, and sustain, innovation and entrepreneurship. Fundamental questions about overlapping and evolving pathways for innovation and regional economic growth remain obscure.

A more nuanced understanding of the sources of spatial and temporal variation within a regional economy is possible by combining data from a variety of electronic sources to gain a more complete picture of constituent organizations, their relationships and the ways in which inventors and entrepreneurs transverse institutional and geographic space. Digital data sources easily permit the construction of detailed firm records useful to studying regional industrial and institutional dynamics. This paper describes our efforts to combine and integrate public and private data sources to create detailed time-series micro-level data. Our framework is designed to study the development of the region by studying extinct, existing and emergent firms and institutions, triangulating data to gain a more complete understanding of the development of firms. The data are organized in a relational database that can support multiple forms of analysis, including mapping and network visualization, statistical modeling or more qualitative approaches. Focusing in-depth on one region removes unobserved heterogeneity that falls to the error term in cross-sectional intra-regional comparisons and permits an examination of the dynamics of innovative economies, specifically how organizations and institutions work together, adapt and improvise to define a functioning regional economic and social system. By recognizing sources of variation within a regional economy, we not only gain a more sophisticated framework for identifying concurrent pathways to nurture the development of innovative firms but also a better understanding of why certain regional attributes predominate and contribute to regional advantage. As such, we are in a better position to situ-

* Corresponding author.
*E-mail address:* maryann.feldman@gmail.com (M. Feldman).

http://dx.doi.org/10.1016/j.respol.2015.01.015
0048-7333/Published by Elsevier B.V.

ate economic actors within an institutional environment, but also examine how their actions and interactions contribute to institutional transformation over time.

The primary purpose of this paper is to describe our approach for combining and creating digital data to study regional economies and industries. Our illustrative project integrates annual data from third-party data sources and tools into a digital infrastructure that is organized around both established and entrepreneurial firms in North Carolina's Research Triangle region. This project extends efforts to bring new electronic sources of private sector data and new analytic tools to the study of dynamic regional economies (Feldman et al., 2012). We describe our data collection efforts that define the region by examining the location of firms and tracking their development and program participation through annual events that are collected from a variety of sources. In partnership with UNC's Renaissance Computing Institute (RENCI) we have designed and implemented a relational database organized around records for over 4200 technology-intensive firms.

Our database design is organized around three components that are discussed in turn. The first is a relational database that tracks firms over time, using a variety of third party data sources. The second component is a data archive that preserves and catalogues studies that collected survey data and other single occurrence variables so that other scholars can re-examine these from their disciplinary perspective. The third component adds context through a digital archive of public documents, reports, and oral histories. Combining these elements, we create a time series data platform linking individuals, firms and institutions while providing historical context to their interactions. In the process, we have created a digital template that would enable scholars to conduct data intensive studies to address fundamental questions about innovation, regional growth and economic development. Though our particular research focus is the Research Triangle region located around North Carolina's Research Triangle Park, in this paper we describe a replicable methodology. We offer our project as an example of how digital data sources can be integrated to better understand regional economies.

## 2. Defining regions by studying firms[1]

The availability of new digital third party data sources calibrated to individual firm addresses provides a means to define a region. After all, places like Silicon Valley or Route 128 are not easily identified on any map; however, they are well known as organizing platforms for innovative economic actors and activity. In practice, innovative locations are defined by the location of firms. Silicon Valley, the prototype for an innovative region, is so named because it was the epicenter for the silicon-based semiconductor industry. Personal computer component makers followed semiconductors, subsequently followed by networking and Internet companies–all creating a revolution in electronics miniaturization and computers, fueled by entrepreneurial firms (Lécuyer, 2008). The local community adopted the name Silicon Valley, which was important to establishing regional identity. Geographically Silicon Valley now encompasses all of the Santa Clara Valley and the southern East Bay. As with any desirable real estate, the boundaries have expanded over time, following yet also influencing the location of entrepreneurial firms, who at times cannot afford prime real estate but want to identify with, and be part of, Silicon Valley.

The Research Triangle, our study region, offers another illustrative example. Research Triangle can refer to the Research Triangle Park (RTP), a 7000-acre industrial park with its own dedicated zip code. But the RTP is a 1960 style low-density research campus of mostly large multinational firms, which until recently had real estate covenants that precluded entrepreneurial firms. Instead, most of the entrepreneurial activity in the region is located in the contiguous communities adjacent to the Park-places with names like Morrisville, Cary and Apex that are not well recognized on their own but are integral locations for economic activity. Often these communities are counted as part of the metropolitan areas of Raleigh or Durham, reflecting the tradeoff between industry detail and geographic specificity. Often finer geographic data are suppressed to maintain confidentiality as required for government data collections. Additionally, regions can be defined by government entities. For example, the state of North Carolina considers the Research Triangle Region to be a 13 county planning region that stretches to the Virginia border and provides data aggregated over this large and diverse region. In reality, there are important synergies between the Research Triangle and the cities of Greensboro and Winston Salem and firms gravitate towards those locations, which are outside the state-defined region. In addition, the larger 13 county region masks the micro-geography, as firms in specific sectors may agglomerate in closer proximity that is dilute when considered against activity in the larger region.

As these examples and other illustrate, we study innovative and entrepreneurial regions that are socially constructed in so far as they do not conform neatly to rigid political boundaries or government statistical units (Amin, 1999). In reality, the functional boundaries of regional economies are reinforced instead by the location of prominent institutions, often universities, government labs or large successful firms, and are influenced by existing transportation routes and land use patterns—factors that in turn influence and reflect firm location (Audretsch et al., 2005; Bathelt et al., 2010). The definition of region is also fluid and expands over time due to idiosyncratic and serendipitous economic and social events (Feldman et al., 2012). Spatial patterns follow a logic that motivates firms to locate near others with similar products, and markets, and close to employees with the requisite workforce skills. The location patterns of related firms define the region (Markusen, 1994). The definition of the region evolves as more firms form and grow, and are attracted to the region.

Examining federally-defined geographic units such as metropolitan areas or counties can also mask cross-border activity and often lead jurisdictions to act as though they are in competition when they could gain from collaboration (Bartik, 1991). Equally important, the geographic co-location of innovative, creative firms in small places such as multi-tenant buildings, neighborhoods, or industrial parks is often invisible when data are available only for larger, administrative units (Sassen, 1989). Firm address level data are important to understanding innovative geographies and allow us to define the region of study based on the behavior of firms rather than the other way around: what appears to be an agglomeration at the county level may indeed be several geographically (and often technologically) distinct groups of firms each with different social relationships and underlying patterns of development (Kohlhase and Ju, 2007).

The technology-intensive firms driving regional economies are themselves fluid and difficult to classify using standard industrial classification schemes. As firms struggle to survive, they often pivot to modify their products or services, but there is no incentive and limited opportunity to update their standard industrial classifications—the main mechanism primarily used to understand industrial activity. New industrial activities, such as 'clean tech' or even optical science defy more standard classifications (Lane, 2011; Clark, 2014). Therefore, by using keywords or combinations of phrases or membership organizations we are able to better assess a broad range of activities and technology applications (Feldman and Lendel, 2010). Both patents and product announcements provide

[1] There is a long academic tradition of studying regions by examining their constituent firms s advocated by Markusen (1994).

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

additional information on the technology and market orientation of firms. Understanding forward-moving industrial activity requires classification schema to be fluid and malleable; based on text mining of patents and product announcements or algorithmic programming to define relational attributes. Static classification schemes will never provide a full understanding of the emerging technologies that have the greatest promise for building new industries and setting regions on a new growth trajectory (Feldman, 2012).

When it comes to firm genesis, the literature has tended to focus either on firms emanating from either research institutions (Shane, 2004; Rothaermel et al., 2007) or being spawned from existing incumbent firms (Klepper 2001, 2002; Klepper and Sleeper, 2005; Klepper and Thompson, 2010; Burton et al., 2002), and has focused on single transactions, such as venture capital investments (Stuart and Sorenson, 2003; Audretsch and Keilbach, 2004), government investments, such as the Small Business Innovation Research (SBIR) program or participation in incubator and accelerator programs (Siegel et al., 2003). Typically, econometric analyses evaluate individual programs, treating them as if they were the only intervention for the entrepreneurial firm. In reality, entrepreneurial firms move from one program to another, and it is the combination of interventions that affect ultimate success. Studies frequently debate the impact of different funding types without considering their interactions and the totality of resources required for the successful launch of an entrepreneurial venture. What is needed therefore is a framework for studying one region, over time and in depth, in order to examine how the various developmental pathways develop and co-evolve over time.

## 3. Circling the triangle

Our project contributes to this endeavor by focusing on North Carolina's Research Triangle region. While the history of the Research Triangle Park (RTP) is well known (Link, 1995, 2002), less is known about the entrepreneurial ecosystem that developed around the park. Our study region offers the advantage of being the spatial scientists' ideal featureless plain when the RTP started in 1958. The region, where the soil had become nutritionally poor from over-farming, was described as dominated by "scrub pines" and "possums" and three universities, the University of North Carolina, North Carolina State University and Duke University, defined the points of the triangle (Wilson, 1999). Our data collection efforts benefit from a punctuated beginning that provides a clear genesis date from which to begin data collection. Still this is not a requirement for replication, but rather an added convenience for us in determining our historical starting point.

At the center of this project is a unique relational database that was initially curated by Dr. William (Bill) F. Little, a longtime faculty member in Chemistry at the University of North Carolina at Chapel Hill, and former Vice President of the UNC system. Bill Little began his career at UNC in 1955 just as North Carolina's Research Triangle Park was getting organized and was a member of the inaugural board of the Research Triangle Park Foundation, an affiliation he maintained until his death in 2009. Unconvinced by a reporter's offhand comment about a perceived lack of entrepreneurial potential in the region, Bill Little commissioned a study of technology-intensive entrepreneurial start-up firms in 1990. His initial list contained 117 firms and included a number of early 'home-grown' entrepreneurial successes, such as Troxler Electronic Laboratories (founded in 1958 by William Troxler in the basement of his Raleigh home), SAS (founded in 1976 by James Goodnight, then a professor at North Carolina State University), and Quintiles (founded in 1982 by Dennis Gillings, then a professor at the University of North Carolina, Chapel Hill). By 2005, Dr. Little had grown the database to 1800 entrepreneurial firms, most of which he identified from regional newspaper articles and through personal correspondence with company founders.

Since 2006, we have significantly expanded the Bill Little dataset by vetting and verifying the original data, triangulating against other data sources, adding more recent new firm startups, and including firms in software, gaming and other sectors that were excluded in initial rounds of data collection.[2] In some respects, our data collection efforts were slowed by the existence of earlier collected data as we had to interpret why certain firms were initially included in the database and spend time debating their removal. As one example, Bill Little choose to include the regional offices of Eastern Airlines on his list, a once prominent airlines that shut down operations in the early 1990s. Still, by going through this additional review, we ultimately clarified our criteria for firm inclusion and also had an opportunity to uncover details on now dormant entrepreneurial and anchor firms that might not have been so easily captured without Bill Little's earlier research contribution.

As a first step, we verified information from the original Bill Little list through a mix of newsprint and online sources, as well as through company and membership directories gathered from prominent support organizations in North Carolina, including the North Carolina Biotechnology Center, Research Triangle Foundation, the Microelectronics Center of North Carolina, and records from the university technology transfer offices, incubators and technical assistance programs. In expanding this data source, our objective is to capture, as much as possible, the full universe of entrepreneurial firms in the region. While Bill Little's original data collection was done primarily by hand and using print materials, the availability of electronic data has greatly accelerated our own data collection efforts. As of December 2013, our updated database includes over 4200 firm records, with 3800 classified as technology-intensive entrepreneurial start-ups, more than doubling Bill Little's initial pool. The following sections describe the organization of the firm database, the data sources used and the founder data we have collected.

## 4. The firm database

We initially created the database in Microsoft Access, but transferred it to SQL as complexity and data sources increased. With technical assistance from UNC-Chapel Hill's Renaissance Computing Institute, we have recently migrated to a more user-friendly web-based infrastructure that better supports simultaneous access.

As captured in Fig. 1, the basic organizing unit of analysis for our relational database is the individual firm, specifically defined here by establishment address location.[3] Establishments are classified under two broad headings: first, large multi-jurisdiction or multinational firms with establishments located within the boundaries of the Research Triangle Park or that have an establishment in the surrounding region and second, entrepreneurial establishments that were started within the region. To track the large multinationals, we use the RTP Directory, which has been published annually

[2] When asked why these firms were omitted from the original data collection Bill Little reported that he was simply more interested in firms related to chemistry and the life sciences and did not think that the software and gaming companies would have much staying power or growth potential. In consultation with the council on entrepreneurial development we were able to find names and addresses for these firms. We then systematically set out to build a comprehensive data set, verifying information on firms from at least 3 different sources before including them in the database.

[3] We differentiate between the firms' original address and current location – using both for different analysis. Annual updates to addresses, starting in 1991 are available through NETS, which tracks establishment location. Most entrepreneurial firms are single establishments initially and the extent to which they expand in the region is one research question currently under investigation.

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

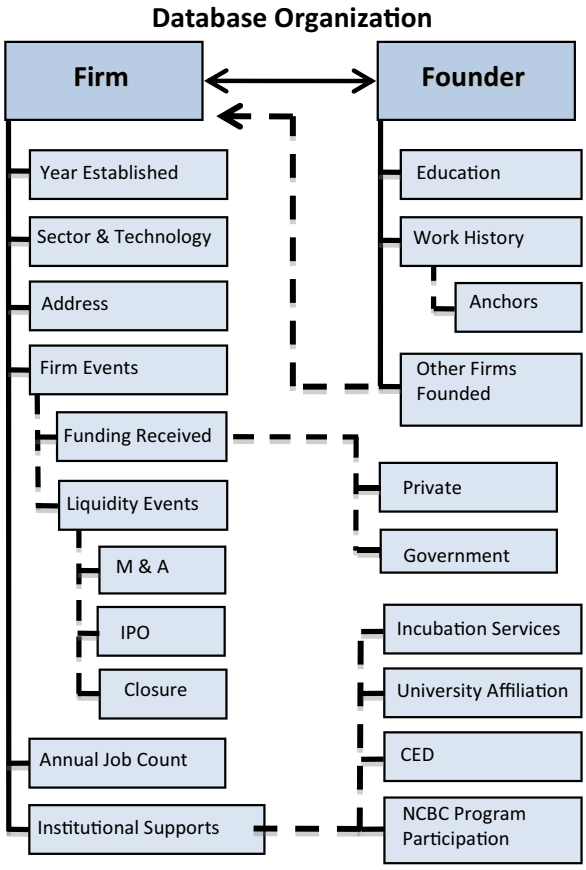


**Fig. 1.** Structure of the database.

since 1968, augmented with the Triangle Business Journal's Book of Lists, which catalogues regional economic activity every year starting in 1992.

For entrepreneurial firms our objective is to capture technology-intensive establishments that are scalable and have the potential to sell goods and services outside the region. We identify new startups through partnerships with local incubators, membership associations, the region's three major research universities, government labs such as the National Institutes of Environmental Health and the Environmental Protection Agency, and other organizations such as the Research Triangle Institute. We also draw heavily from sector specific support organizations, which maintain detailed and searchable directories of entrepreneurial firms, including North Carolina Biotechnology Center and the Microelectronics Center of North Carolina. We also use newspaper clippings and web scraping to track new firm formation and development. Many mentions of new firms occur when the idea is still in the formative stage. Other firms are identified when they apply for patents or receive financing, beginning with little fanfare or public notice. Our strategy therefore has been to triangulate and validate information, noting discrepancies to understand potential bias in different sources.

When determining whether a firm should be included in our database, we look for at least three mentions in the various sources we track (see Table 1). As one source, we rely on records from the North Carolina Secretary of State, which requires all firms conducting business in the state to register and pay an annual fee. It is often difficult to pinpoint the exact startup date for an entrepreneurial firm, as different dates appear to promulgate for different purposes. Similar to individuals, firms want to be older when they are getting started as a means to establish legitimacy and report a later date when they want to appear fresh and innovative. The date of incorporation or registration to conduct business is therefore used as the recorded date when a project or idea becomes a firm and is included in the database. The Secretary of State also permits a tracking of firm deaths, given that firms registration lapse when the firm ceases to exist.[4] Indeed, over 40% of the entrepreneurs in the database are no longer in existence, having been acquired, merged or gone out of business. Thus, these data provide a resource to study the complexity of regional industrial dynamics while avoiding a common truncation of the phenomenon by studying only a subset of firms that have survived, thereby biasing the analysis (Lichtenstein et al., 2007).

**Table 1**
Annual data sources, matched to firm name, listed alphabetically.

| Primary data source | Type of information |
|---|---|
| Council for Entrepreneurial Development (CED) | Program participation Institutional support |
| Delphion patent data | Granted patents Patent applications |
| Innovaro Medical Device Licensing | Licensing agreements |
| Microelectronics Center of North Carolina | Institutional support Archived newspaper clippings Annual reports Meeting minutes |
| National Establishment Time Series database (NETS) | Firm location Employment Sales growth Parent company affiliations |
| National Venture Capital Association | Venture financing |
| North Carolina Biotechnology Center (NCBC) | Firm description Product development; clinical trials Institutional support |
| One N.C. Small Business Program | State matching institutional support |
| Quarterly Census of Employment and Wages, North Carolina | Establishment level employment Average wages |
| Small Business Association—SBIR/STTR program | Federal SBIR funding |
| S&P Capital IQ | Financial performance Investment |
| ThomasNet.com | New product introductions |
| U.S. Food and Drug Administration | Status of FDA approval |
| U.S. Securities and Exchange Commission | Filings for IPO |

### 4.1. Firm development data sources

Once a decision is made to include a firm in the database, detailed information about the firm and its organizational, financial and technological development is gathered from company websites, annual reports, press releases, newspaper articles and social networking sites. Using firms as our common frame of reference allows us to evaluate the total sum of resources that entrepreneurial firms use in their development and the ways in which interactions between resources shape firm development. While many of the financial and institutional data sources we use are public, our core contribution is bringing these together into a single repository for analysis.

Company webpages and press coverage are often a good source for this kind of information, but require special handing to place into a form useable for a database. There are a plethora of third party data sources that we use to track firm development overtime. These data capture annual events such as inventive activity, new product announcements or annual financial performance and employment. We also use third party data sources available through the university library system and research labs. These data are often used for business analysis and their use is often

[4] While the secretary of state data are a reliable sources for vetting firm information proven less useful for identifying new firms: there are more than 400,000 entities active at any time, with all business activity in the state reporting.

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

limited to projects in finance in business schools. However, the third party datasets that we incorporate also have great utility for examining firm dynamics. For example, patent data are a good source of inventive activity as covered by other papers in this special issue. But other sources include ThomasNet.com which provides information on new product announcements and the National Establishment Time Series for tracking sales and employment growth. The National Establishment Time Series, created by Don Walls, also provides us with information on company location and changes over time and details on name changes, mergers, acquisitions and closures. We supplement these data with annual information about employment (from the Quarterly Census of Employment and Wages (QCEW)[5]); financial performance and investment (NETS; S&P Capital IQ; National Venture Capital Association); patenting activities (Delphion); and product development (U.S. Food and Drug Administration; ThomasNet.com). We also monitor newsfeeds to augment these sources. Fig. 1 and Table 1 provide a more comprehensive overview of the specific types of firm-level development data we capture and their sources of origin.

Third party data offer advantages over collecting original survey data. These sources cover a population of firms and avoid the limitations of self-reporting, and low response rates that make original data collection difficult. Instead of relying on a single data source it is possible to integrate different data sources matched to firm name and address. Our common frame of reference allows us to evaluate how differences in empirical findings, for example related to firm employment, location or venture capital financing, are driven by the choice of data source (Donegan, 2014). Moreover, the sources we rely on are updated continuously, offering an opportunity to gather timely information useful to policy makers.

Given their importance for regional entrepreneurship, our database includes a set of prominent anchor firms, many that are multinational or multijurisdictional establishments that relocated to the region or established a sizeable branch plant. Still, the vast majority of firms in our database are entrepreneurial. For these entrepreneurial ventures in particular, we have also collected detailed information on Initial Public Offerings (IPO), funding from the Small Business Innovation Research (SBIR) program and state programs, notably the North Carolina Biotechnology Center and One NC Small Business Program. We also track participation in programs run by local entrepreneurial support organizations, such as the Council for Entrepreneurial Development and local incubators, accelerators, and angel groups. For each firm, we trace the annual level and sequence of institutional support. We found these institutional partners to be engaged participants and willing to share membership or program participation lists. Similar private and quasi-public entities devoted to innovation, entrepreneurship and economic development exist in other regions and organizations like SSTI.org their presence and role in the regional economy, suggesting again an opportunity for replication.

### 4.2. The founders

We have designed our database to capture detailed career and educational histories for all entrepreneurial founders. Our database allows us to trace numerous entrepreneurial firms in the Research Triangle region back to prominent North Carolina universities via official university technology transfer channels. But this is not the only source of firm genesis. Large multi-jurisdictional corporations in the region have also helped to incubate entrepreneurial talent and skill. By studying firm founders and their career histories we are able to identify the co-existence of these pathways and thus, open up the potential for comparative analysis. Ultimately founder data, combined with other statistical and qualitative sources, enables us to study the construction and evolution of multiple entrepreneurial pathways and in the process, identify distinct and intersecting organizational and institutional influences.

Founder information was gathered from a number of digital and archival sources. Media accounts of new firm formation often include the names and prior organizational affiliations of members of a founding team. In addition, the names of firm founders are available through incorporation documents filed with the North Carolina Secretary of State. Extending from these sources, we rely heavily on social media, namely LinkedIn, news sources, company websites, alumni records and interviews to capture detailed information on founder education and career histories. Where possible, we have tried to include founder information for university or college degrees earned at public, private, large and small universities both within and outside the region. Using past employment information from LinkedIn, we also trace firm founders back to prominent multinational firms within the region, such as Glaxo, Burroughs–Wellcome, Becton Dickinson, and IBM, among others. But equally, we capture past employment at 'home-grown' establishments, including smaller firms like Addrenex, or Icagen, thereby allowing us to identify second-generation entrepreneurs.

We are also in the process of conducting interviews and oral histories with firm founders and top-ranked executives to understand more about the regional entrepreneurial experience. We treat these qualitative data sources as especially valuable to the study of regional innovation and as a resource for contextualizing statistical patterns and econometric relationships and also for supporting a mixed-method research design. With that in mind, we have initiated collaboration with the Southern Oral History Project (http://sohp.org/), a rich digital repository stored at UNC-Chapel Hill containing over 4000 southern oral histories. Staff from UNC's Southern Oral History Project provide training for graduate students and faculty on our project, in exchange for contributions to the Collection of taped and transcribed oral histories with entrepreneurial founders.

### 4.3. Historical context

In the process of compiling firm-level and statistical data, we have also amassed a large and growing collection of archival materials from quasi-public, private and non-profit organizations and support institutions located in the region. These documents range from organizational directories and annual reports, to meeting minutes for committees created to design technology support policies and programs. They also include the names of institutional board of directors, as well as historic information on firms and individuals receiving early programmatic support. These efforts complement an existing archive at the UNC Library on the Research Triangle Foundation, which contains 88,000 items constituting 125 feet of linear shelf space. Our objective was to broaden this collection to include activity outside the Park, which is the focus of the Foundation and to also create a digital archive.

Far too often institutional documents and policy reports important to understanding the historical development of the regional economy are in danger of being lost. In some cases these historic documents have been retrieved by what may only be described as dumpster diving as organizations have changed locations or mission resulting in a need to cull earlier collections and paperwork (Feldman and Lowe, 2011b). Some organizations keep updated digital records on firms they support. The North Carolina Biotechnology Center, for example, maintains a rich, continually updated source of annual reports, new releases and programmatic information. Their collection is well cataloged and digitally preserved,

[5] Disaggregated establishment-level data from the quarterly census of employment and wages (QCEW) were provided under special arrangements by the North Carolina Department of Commerce.

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

allowing for ease of use and document retrieval. For the now defunct or under-resourced regional support agencies lacking a formal archival method, annual reports become essential for identifying their earlier contributions to industry and firm development. In other cases, we have collected documents from the personal files of individuals we have interviewed. We intend for these archived data sources to be regularly updated, thereby providing interdisciplinary research teams with a more meaningful contextually grounded research experience and encouraging mixed-method research on this region.

The risk of losing valuable knowledge about the regional economy motivates us to continue to collect and digitize older publications and reports. In turn, these documents support more in-depth and contextual understandings and interpretations of firm and founder-level data. But equally, these archival documents allow us to understand the broader policy environment and influence on firm formation and development.

By incorporating and categorizing information on various regional support programs, we are able to gauge the scope and depth of regional institutional engagement by individual firms.

Still our interest in collecting institutional data is not simply an attempt to understand the institutional impact on firm performance. Our goal is to also situate firms as actors within a changing institutional landscape. In this regard, we recognize firms as more than mere inheritors of a regional institutional environment. They are also active participants influencing and inspiring institutional change over time as they move through a regions' institutional space. Using archival sources therefore allows us to see how the changing institutional landscape of the Research Triangle region is influenced by firm-level strategy.

But equally, we use these sources to identify other actors—including those within and outside established regional institutions—that shape how those institutions respond to new pressures and challenges and in ways that may deviate from the intended impact of intervening firms, including those with considerable power and authority. In this regard, we recognize institutional change as an evolving and mediated process, involving multiple actors and open to negotiation and reinterpretation (Feldman and Lowe, 2011a,b).

## 5. Analytical example 1: the case of GSK

As an illustrative example of our database and its analytical potential, we have used founder employment records compiled from LinkedIn and other sources to trace 144 firms to GlaxoSmithKline (GSK) and its antecedents of Burroughs–Wellcome and Glaxo–Wellcome (Feldman and Lowe, 2014; collectively referred to as Glaxo). This has enabled us to engage in a long standing academic debate over 'spawning' of new entrepreneurial firms from the stock of existing and well-established firms (Chatterji, 2009; Klepper, 2001). Spawns are distinct from firms that spin-out from government laboratories or universities. Spawning instead refers to an employment relationship when the new company founder was previously employed at another private sector entity. Spawns may be either formal or informal spinoffs from a parent company. Thus while the number and type of entrepreneurial startups in a region is an important indicator of economic robustness, the number and type of spawns provides additional information about the connections between large and small firms in the region and the degree to which founders benefit from employment and social ties to other regional firms.

The vast majority of the Glaxo spawns we have identified were established in the wake of acquisitions, mergers and major corporate restructurings, which resulted in both voluntary and involuntary employment termination. Our relational database enables us to track the formation and development of these entrepreneurial spawns overtime, including business type, technology specialization, commercial success, employment growth and level of institutional support, including external financing. Equally, we are able to capture founder employment rank at the Glaxo firms, which in turn affected their level of severance support and access to intellectual property upon departure. Combining these data with founder interviews and archival materials gathered on the history of the Glaxo group, allows us to trace changes in the way GSK and its predecessors supported entrepreneurial development in the region—this is an underexplored area in well-established literature on spawning. In turn, we are better able to capture changing organizational strategies at GSK and its predecessors and explore how they have influenced the types of new business establishments spawned over the years and also affected the industrial structure in the region.

We drew on the database contents to identify Glaxo spawns and then conducted in-depth interviews with a representative sample of firm founders. In addition we held focus groups for different cohorts of founders, based on their years of employment at GSK and its predecessors. Through our analysis, we discovered that a sizeable share of these entrepreneurial spawns established in the wake of the first merger between Glaxo and Burroughs–Wellcome specialized in contract research and clinical trial management organizations rather than de novo drug discovery or new product development. Using archival materials, we determined this coincided with a 1990s strategy shift at Glaxo–Wellcome towards greater outsourcing of research and clinical trial development. In turn, this first wave of service-oriented entrepreneurial spawns contributed greatly to the build out of North Carolina's burgeoning contract research (CRO) industry—today, the Research Triangle region is home to approximately 150 CROs, the largest regional concentration in the world. Our interviews also revealed the importance of statistical expertise at local universities and the presence of firms like SAS that provided sophisticated analytical tools and thus, a competitive edge for newcomers in contract research. Equally, early CRO firms in the region, like Quintiles and Pharmaceutical Product Development (PPD), helped provide a role model and inspired others to follow suit. Still, our analysis suggests the development of this specialized sub-sector of life science was strongly influenced by the practices and strategies of Glaxo–Wellcome as an influential anchor firm, a contribution that the current literature has yet to fully recognize. In fact, executives at Glaxo–Wellcome were especially supportive, encouraging former employees to set up regional CROs with promises of long-term contracts.

Based on our analysis, 55% of entrepreneurial firms created shortly after the merger of Glaxo and Burroughs–Wellcome specialized in some form of contract research or analytics support. In contrast, the vast majority (61%) of new firms created in the wake of the 2000 merger involving Glaxo–Wellcome and SmithKline Beecham specialized in new drug discovery or medical device development—with significantly less (32%) focusing on contract research and related analytics. Some of this shift in focus reflected the establishment of formal licensing agreements with GlaxoSmithKline (GSK) that enabled bench scientists to take their research program and ideas with them to their new firms. Strong internal channels for asset development and out licensing supported this. But equally, regional institutions were far more developed at the time of the second merger and our interviews with firm founders confirm these institutions offered a wider range of support services for technology and business development.

GSK and its predecessors are not alone in creating a lasting regional entrepreneurial imprint. As a result of our continued data collection efforts, we have identified a dozen or so large multi-jurisdictional firms other than those in the Glaxo group that have spawned sizeable numbers of entrepreneurial firms. These include

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

IBM, Nortel, Sony-Ericsson, Red Hat and Becton–Dickinson & Co., to name a few. Spawning by these firms is so essential to the regional entrepreneurial economy that we find virtual parity in the number of entrepreneurial start-ups attributable to the region's three major research universities and those attributable to the top three spawning corporations. While several of these spawn-producing firms experienced periods of downsizing, others like Red Hat and Becton–Dickinson have been operating mostly in high-growth mode thus allowing us to examine additional factors that might shape cross-organizational differences in entrepreneurial spawning. As we consider the potential influence of large anchor firms on industrial and entrepreneurial development in the region, we are currently taking steps to identify entities that are less likely to spawn new establishments. One important example in the Research Triangle is SAS Software, a firm that initially spun-out of North Carolina State University in 1976, however has not resulted in many next generation start-ups. That said SAS still has a regional entrepreneurial influence, which our database allows us to trace through their involvement in key support institutions.

### 5.1. Analytical example 2: the NC biotech center

A further illustration of the analytical contribution of our database involves the North Carolina Biotechnology Center (Biotech Center)—a prominent industry support institution in the Research Triangle established in 1981—and more specifically its efforts to better support agricultural biotechnology in the late 1980s. For this case study, we drew mostly on in-depth interviews and archival documents, though our database also enabled us to capture the extent of institutional support the Biotech Center has provided to agricultural biotechnology firms. The Biotech Center stepped in to support agricultural biotechnology after learning that field experimentation by a large multinational firm could have potentially damaging effects on neighboring farms. Admittedly, the multinational firm in question had hoped its initial interaction with the Biotech Center would result in a public relations campaign to assuage the fears of local farmers and environmental activists. Instead, the Biotech Center responded by formalizing a new institutional channel through which local farming and environmental interests could engage with large pharmaceutical corporations and in ways that forced consensus policy around the regulation of agricultural biotechnology. The result of this exchange was our nation's first state-level field experimentation ordinance: an ordinance explicitly designed to reduce risk and uncertainty for agricultural biotechnology firms—including 'home-grown' entrepreneurial establishments—through the creation of a transparent and legally defensible regulatory apparatus. As this case and others like it from the region suggest, institutional changes initially triggered by large corporate interests can strengthen opportunities for policy development in support of regional entrepreneurship and in ways that are unanticipated by those initiating institutional action.

This example and others speak to the broader policy lessons offered by our research. Many places around the world look to North Carolina's Research Triangle as an example of successful policy-led economic development yet the precise mechanisms and policy levers are not well known. As a result, there is a tendency to downplay the influence of policy interventions and interactions on firm strategy and development. Our database and supporting documents offer a data-driven solution to this challenge by providing an essential resource for tracing the sequence of institutional interventions and facilitating policy events that contribute to and help sustain entrepreneurial development. It also helps to reveal the role of institutional actors in shaping the co-existence of multiple entrepreneurial pathways that support new firm formation. As we have discovered, entrepreneurial firms are just as likely to emerge from university-based technology transfer systems as they are from large multinational corporations that incubate and spawn entrepreneurial talent and skill. Detailed information of this kind will enable practitioners and policy makers in North Carolina to better demonstrate and defend the efficacy of economic development decisions and investments—some of which are under threat in the face of changing state politics and funding priorities. Equally, it will allow development practitioners to identify opportunities for leveraging scarce regional resources by exploring synergies with federal, state and regional agencies.

## 6. Challenges and considerations for replication

Our Research Triangle study provides an initial test case for compiling and collecting spatial and temporal data at the firm, institutional and individual actor levels. In this paper, we have presented our approach for integrating and preserving data from prior studies, incorporating archival and electronic data sources and navigating new social media. In this regard, our project creates a data-intensive loosely coupled framework and replicable methodology for broadly analyzing economic and social dynamics in other regions around the world. Still, our on-going data collection effort has yielded transferable lessons, and some persistent challenges, which we continue to work through and learn from. What insights and advice might we therefore offer other researchers that are seeking to compile similar comprehensive data sets in other regional economies?

First, community building is essential to our on-going data collection effort. Countless individuals and organizations have contributed to this exercise, providing essential data, historical context and interpretation. This project has benefited tremendously from partnership with UNC's Renaissance Computing Institute (RENCI), a collaborative institute between the three Triangle universities that creates software tools, visualizations, and designs data management systems. RENCI has been a technical resource and a true partner, building and maintaining a relational database, which is described in greater detail in the next section. This work has also benefited from data sharing agreements with local organizations, including the Council on Entrepreneurial Development, First Flight Venture Center and the North Carolina Biotech Center.

We have intentionally designed this project to prioritize student involvement and education. Since 2006 we have involved approximately 25 graduate and undergraduate students, three-quarters of whom are women, minorities or international students. In an effort to stimulate additional student interest in this project and data source, we also use class instruction time selectively to feature the database.[6] We have worked with professional masters students from City and Regional Planning and the School of Library and Information Science, who have used the database for coursework, independent study or research projects and in the process made suggestions for how to improve data collection and organization. Finally, this project has benefited from a small army of undergraduates, who have worked under a variety of different compensation schemes, including additional data collection for independent studies and employment paid from a variety of sources. Undergrads have told us that experience researching companies and building the database was valuable in subsequently securing employment. We have also engaged with a number of outside researchers that have requested data access for their own research purposes.

[6] Ted Zoller and Lisa Goble have used these data in their completed doctoral dissertations and three other dissertation projects are in the works. Gil Avnemilch was a force in the beginning of the project and Max Peter Menzel worked on visualizing networks with the data.

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

But with this extended community comes a nested set of data governance challenges. First is the need to clearly specify rules and routines for those engaged in data entry and cleaning about what details to add, modify or delete from our evolving relational database. This has been a work-in-progress and hence a learning opportunity for us and other project participants. In particular, as a result of early missteps that risked data loss or inconsistent reporting, we have come to recognize the value of time stamping all database entries and edits to allow for greater oversight and accountability. Equally, data governance issues have arisen in relation to recent requests for data access by individuals outside our core research community. These requests required us to specify formal rules to make data available, especially to non-contributing users. In response, we developed a data sharing agreement, customized to each data request, which outlines our expectations for authorship, attribution or data sharing reciprocity. We currently have two contracts on file, both with graduate students requesting parts of our dataset for dissertation or thesis research.

A second major goal (and thus learning opportunity) for us involves the creation of an accessible infrastructure for regional policy analysis and theory building. Our vision is a flexible cyber-enabled resource that policy makers, practitioners and scholars can utilize for identifying regional trends and challenges to technology development and entrepreneurship. Unfortunately, our present data storage system is not fully equipped to support this use, namely because it cannot selectively suppress data at the individual establishment or firm level. As noted earlier, our relational database includes information at the firm or establishment-level on annual employment, venture financing and grant awards, some of which was shared with us with an expectation or contractual requirement that we keep individual firm names and award amounts confidential. This creates the need for a layered data infrastructure that can accommodate multiple users with different levels of data access and thus at varying levels of aggregation. This problem is one that we are currently working to resolve with assistance from our partner RENCI and with funding support from the National Science Foundation. What we propose is a data management and integration infrastructure based upon leading edge and widely used research software systems that can support data integration, analysis and visualization capabilities. Attempts to replicate our data collective effort elsewhere would therefore benefit from a concurrent commitment to infrastructure development.

Additionally, we continue to struggle with the time consuming tasks of matching data from different, independent sources that do not share a numerical firm-specific identifier. This is a common disambiguation challenge shared by numerous scholars that seek to combine multiple, large sized data sets. To paraphrase one computer scientist working on our project, this is essentially a billion dollar computational problem. While many scholars studying firms and establishments have undertaken similar matching processes, few have documented their techniques, hindering our collective ability to arrive at a "best" approach to this problem (see Donegan, 2014). For our part, we have used a semi-automated strategy. We first rely on a SAS program to match firm names from our database to names from other datasets (see Fuchs et al., in this issue for greater detail on the disambiguation problem, and Donegan, 2014 for further description of the specific SAS command we use). The program generates a list of the top 20 potential matches for each firm name in our database. Members of our team then review these lists by comparing the names and addresses from our database to names and addresses returned from the other database. Records with the same name and address in both databases are considered a match. In some cases, such as when firm names match but addresses are different, the team extends their search to include other sources, such as Secretary of State incorporation documents or firm websites to determine if a match can be made. Additionally, two members of the team are assigned identical lists of potential matches in an effort to minimize problems of subjectivity (i.e., all matches are double-coded). While this phase is data and labor intensive, once unique identification codes are assigned, annual updating becomes straightforward, with the need to only research new firms or firms that have changed names or ownership. Moving forward, our hope is that other research communities will see value in developing and openly sharing code and processes for less labor-intensive disambiguation solutions, with the goal of expediting data matching efforts.

On a final note, there is the issue of time and resource commitment and whether this justifies or permits replication. We acknowledge our own efforts have stretched over a 6 year time span and have required considerable financial resources to support a team of graduate and undergraduate researchers. Still, much of this time was spent vetting earlier data, exploring public data sources, determining criteria for firm and founder inclusion, testing various matching methods and identifying the appropriate cyber-infrastructure needed to house our database and support documents. The time span also reflects the wide range of industries included in our database, and the "start-up" time of building a physical database and research process from the ground up. Much has been learned by our team in the process and we believe our insights could benefit other research groups and ultimately help expedite attempts to replicate similar data collection and storage efforts elsewhere.

## 7. Concluding reflections

By outlining our approach in this paper, a key objective is to offer a transferable framework for analyzing regional dynamics in other locations. Today, such an effort would not have been possible without digital data and an expandable relational database. Rather than rely solely on interviews, our combined approach outlined here enables a temporal snapshot based on the confluence of data. This allows for better identification of candidates for interviews as well more targeted questions during interviews that fill in details or resolve seeming contradictions. For example, database queries enable the identification of founders with prior employment connections to specific firms or anchor institutions and therefore enables us to better study entrepreneurial cohort effects, like those described above for firms with established links to GSK and its predecessors.

We therefore offer the Research Triangle Region as a test case for compiling and collecting data at both the firm and institutional levels that would be transferable to other regions. Our approach offers a data-driven resource that scholars, policy makers and practitioners can use to trace the sequence of interventions and facilitating events that contribute to and help sustain entrepreneurial development. It also helps to reveal the existence of multiple entrepreneurial pathways that support new firm formation. Ultimately, detailed information of this kind will enable researchers and practitioners to better demonstrate and defend the efficacy of regional economic development decisions and investments and further enhance understanding of and coordination among multiple interventions that contribute to the region's entrepreneurial eco-system. This region benefits from a punctuated beginning date. Other research designs may focus on other transformative events and conditions. The critical advance we offer is greater understanding of regional and industrial dynamics.

Sophisticated integrated analysis of social and economic processes has previously been hindered by the lack of meaningful and accessible data. Legacy federal economic data products rarely meet data user needs because the traditional mission of federal economic statistical agencies is to support federal macroeconomic policy and

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015

guide the distribution of federal funds to political jurisdictions with particular economic characteristics, such as high unemployment. Long-standing forms of economic data also do not allow analysts to view the behavior and movement of actors over time. Furthermore, attempts by the Census Regional Data Center to provide access to micro-data at the firm and individual level have been stalled due to disputes with the Internal Revenue Service and analysis is limited by Federal Section 13 privacy concerns.

Additionally, traditional datasets and industrial classification schemes do not allow for consideration of the relationship among diverse organizations across space. This includes the ways established corporate anchors spawn small, entrepreneurial firms, some encouraging their formation while others enforcing non-compete agreements (Chatterji, 2009). Equally the relationship between firms and institutions are difficult to track; this includes linkages to universities, trade associations, business services, and other quasi-government entities that research has demonstrated as essential to innovative activity (Asheim and Coenen, 2005; Geels, 2004; Lowe and Gertler, 2009; Owen-Smith and Powell, 2006; Wolfe and Gertler, 2002). These institutions are often the glue that holds a regional economy together providing networks and support services that are the foundation for economic vitality.

In response to these limitations, the data we have been collecting over the past six years is drawn from accessible and frequently updated records, scraped from the web, pulled from voluminous documents through text analysis, found on open data platforms, purchased from third parties by our university libraries, and integrated with other digital data sets. Furthermore, our data permit geocoding at the establishment level, thereby allowing researchers to analyze these micro-data within unique self-defined economic boundaries.

## Acknowledgements

Support for this project is from UNC's Office of Economic and Business Development and the Odum Institute for Social Science Research. The Science of Science Policy Program at the National Science Foundation has provided funding.

## References

Amin, A., 1999. An institutionalist perspective on regional economic development. Int. J. Urban Reg. Res. 23 (2), 365–378.
Asheim, B., Coenen, L., 2005. Knowledge bases and regional innovation systems: comparing nordic clusters. Res. Policy 34 (8), 1173–1190.
Audretsch, D.B., Keilbach, M., 2004. Entrepreneurship Capital: Determinants and Impact. CEPR Discussion Papers 4905, C.E.P.R. Discussion Papers.
Audretsch, D.B., Lehmann, E.E., Warning, S., 2005. University spillovers and new firm location. Res. Policy 34 (7), 1113–1122.
Bartik, T.J., 1991. Who Benefits from State and Local Economic Development Policies? Upjohn Press, Kalamazoo, MI.
Bathelt, H., Kogler, D.F., Munro, A.K., 2010. A knowledge-based typology of university spin-offs in the context of regional economic development. Technovation 30 (9), 519–532.
Burton, M.D., Sorensen, J.B., Beckman, C.M., 2002. Coming from good stock: career histories and new venture formation. In: Lounsbury, M., Ventresca, M.J. (Eds.), Research in the Sociology of Organizations, vol. 19. Elsevier, Waltham, MA.
Chatterji, A., 2009. Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. Strat. Manage. J. 30 (2), 185–206.
Clark, J., 2014. Hidden in plain sight: the North American optics and photonics industry. In: Clark, J., Vanchan, V., Bryson, J.B. (Eds.), The Handbook of Manufacturing Industries in the World Economy. Edward Elgar Publishing.
Donegan, M., 2014. Inside the Triangle: Does Database Selection Alter our Understanding of Urban Industrial Systems? Presented at the Workshop on Big Data and Urban Informatics. University of Illinois at Chicago, IL, Chicago https://dl.dropboxusercontent.com/u/35674979/CFP/proceedings/bduic2014_submission_65.pdf
Feldman, M.P., Lendel, I., 2010. Under the lens: the geography of optical science as an emerging industry. Econ. Geogr. 86 (2), 147–171.
Feldman, M.P., Lowe, N., 2011a. Restructuring for resilience. Innov. Technol. Govern. Global. 6 (1), 129–146.
Feldman, M.P., Lowe, N., 2011b. Industrial Genesis–Southern Style. Presentation to the Association of American Geographers (AAG), Seattle, WA, April 16.
Feldman, M.P., et al. 2012 Innovative Data Sources for Regional Economic Analysis. ebook Washington, DC 2012.
Feldman, M.P., Lowe, N., 2014. Firm Strategy and the Wealth of Regions. Mimeo. University of North Carolina, Chapel Hill.
Geels, F., 2004. From Sectoral Systems of Innovation to Socio-Technical Systems: Insights About Dynamics and Change from Sociology and Institutional Theory. Res. Policy 33 (6-7), 897–920.
Klepper, S., 2001. Employee startups in high-tech industries. Ind. Corp. Change 10 (3), 639–674.
Klepper, S., 2002. The capabilities of new firms and the evolution of the U.S. automobile industry. Ind. Corp. Change 11 (4), 645–666.
Klepper, S., Thompson, P., 2010. Disagreements and intra-industry spinoffs. Int. J. Ind. Organiz. 28 (5), 526–538.
Klepper, S., Sleeper, S., 2005. Entry by spin-offs. Manage. Sci. 51 (8), 1291–1306.
Kohlhase, J.E., Ju, X., 2007. Firm location in a polycentric city: the effects of taxes and agglomeration economies on location decisions. Environ. Plann. C 25 (5), 671.
Lane, E.L., 2011. Clean Tech Intellectual Property: Eco-marks, Green Patents, and Green Innovation. Oxford University Press.
Lécuyer, C., 2008. Making Silicon Valley: Innovation and the Growth of High Tech, 1930–1970. MIT Press, Cambridge, MA.
Lichtenstein, B.B., Carter, N., Dooley, M., Gartner, K., 2007. Exploring the temporal dynamics of organizational emergence. J. Bus. Venturing 22, 236–261.
Link, A.N., 1995. A Generosity of Spirit: The Early History of the Research Triangle Park. Research Triangle Foundation of North Carolina, Research Triangle Park, NC.
Link, A.N., 2002. From Seed to Harvest: The Growth of the Research Triangle Park. Research Triangle Foundation of North Carolina, Research Triangle Park, NC.
Lowe, N., Gertler, M., 2009. Building on diversity: institutional foundations of hybrid strategies in Toronto's life sciences. Reg. Stud. 43 (4), 589–603.
Markusen, A., 1994. Studying regions by studying firms. Prof. Geogr. 46, 477–490.
Owen-Smith, J., Powell, W., 2006. Accounting for emergence and novelty in Boston and Bay Area biotechnology. In: Braunerhjelm, P., Feldman, M.P. (Eds.), Cluster Genesis: Technology-Based Industrial Development. Oxford University Press, Oxford.
Rothaermel, F.T., Agung, S.D., Jiang, L., 2007. University entrepreneurship: a taxonomy of the literature. Ind. Corp. Change 16 (4), 691–791.
Sassen, S., 1989. New York city's informal economy. In: Portes, A., Castells, M., Benton, L.A. (Eds.), In, The Informal Economy: Studies in Advanced and Less Developed Countries. The Johns Hopkins University Press, Baltimore, MD.
Shane, S., 2004. Academic Entrepreneurship: University spinoffs and wealth creation. Edward Elgar, Cheltenham.
Siegel, D.S., Westhead, P., Wright, M., 2003. Assessing the impact of science parks on the research productivity of firms: exploratory evidence from the United Kingdom. Int. J. Ind. Organiz. 9, 135–1369.
Stuart, T.E., Sorenson, O., 2015. Liquidity events and the geographic distribution of entrepreneurial activity. Adm. Sci. Q. 48 (2), 175–201.
Wilson, J., 1999. North Carolina's Research Triangle Park: An Investment in the Future. RTP Foundation, Raleigh, RTF Video.
Wolfe, D., Gertler, M., 2002. Innovation and social learning: an introduction. In: Gertler, M., Wolfe, D. (Eds.), Innovation and Social Learning: Institutional Adaptation in an Era of Technological Change. Palgrave Macmillan, New York.

Please cite this article in press as: Feldman, M., Lowe, N., Triangulating regional economies: Realizing the promise of digital data. Res. Policy (2015), http://dx.doi.org/10.1016/j.respol.2015.01.015