FUNG INSTITUTE
for ENGINEERING LEADERSHIP
UC BERKELEY COLLEGE OF ENGINEERING

MobilityMapper: Visualizations of Inventor Mobility, 1975 - 2010

Guan-Cheng Li
Department of Electrical Engineering and Computer Sciences
UC Berkeley
guanchengli@eecs.berkeley.edu

Lee Fleming
Fung Institute for Engineering Leadership
UC Berkeley
lfleming@ieor.berkeley.edu

College of Engineering
University of California, Berkeley

Fung Technical Report No. 2013.06.18
http://www.funginstitute.berkeley.edu/sites/default/files/Mobility_Mapper.pdf

June 18, 2013

The Coleman Fung Institute for Engineering Leadership, launched in January 2010, prepares engineers and scientists – from students to seasoned professionals – with the multidisciplinary skills to lead enterprises of all scales, in industry, government and the nonprofit sector.

Headquartered in UC Berkeley's College of Engineering and built on the foundation laid by the College's Center for Entrepreneurship & Technology, the Fung Institute combines leadership coursework in technology innovation and management with intensive study in an area of industry specialization. This integrated knowledge cultivates leaders who can make insightful decisions with the confidence that comes from a synthesized understanding of technological, marketplace and operational implications.

We wish to thank the National Science Foundation for supporting this work, with grant #1064182, as well as the Coleman Fung Institute for Engineering Leadership. All errors remain ours. Code is available from the authors.

Abstract: Knowledge industries and the professionals that create and run them have become increasingly important to the economy. Hence, it is helpful to understand 1) where such professionals live, 2) their geographical mobility, and 3) when and why they choose to move. To facilitate such understanding, and visualize econometric results, we build an interactive and dynamic visualization of inter-state mobility of US inventors, from 1975-2010, based on our disambiguation of the US patent database. Focusing on Michigan, for example, one can see a brain drain out of the state after 1985, caused by an inadvertent enabling of non-compete enforcement. The application is at http://funglab.berkeley.edu/mobility/.

Introduction and Motivation

A growing line of research has used patent records to study inventor mobility, often in the study of regional economic and technology dynamics (Almeida and Kogut 1999, Agrawal et al. 2006; Breschi and Lissoni 2009, Marx et al. 2009). Most of this research has relied on manual or ad hoc disambiguation and focused on intra-regional mobility. Automated disambiguation of entire patent corpora enables study of individual mobility of whole populations, across all regions. Visualization tools allow us to "see" econometric results and communicate the essence of the phenomenon to a non-statistical audience.

To illustrate this, we use results from Marx et al. (2012) that demonstrate a brain drain from states that enforce noncompetes to those that do not. The argument is that inventors, especially those with greater human and social capital, will seek opportunities in other regions, when their local opportunities are limited, because they are constrained from working for their employers' competitors within their current region. The research exploited a natural experiment, when the Michigan legislature inadvertently enabled the prosecution of noncompetes beginning in 1985. The identification relied on a differences-in-differences methodology, which compared emigration from Michigan to emigration from other control states that prohibited enforcement of noncompetes over the entire time period of study, from 1975-1996. Marx et al. also provide corroborating cross-sectional evidence for all U.S. states, from 1975-2005.

Figures 1 and 2 illustrate a noticeable increase in emigration from Michigan, comparing 1982 to 1987. Figure 3 illustrates how this emigration was not balanced by immigration. Figure 4 illustrates emigration into California at the height of the technology boom in 2000. These figures are screen shots from the interactive MobilityMapper tool. The tool is capable of "movies" that illustrate mobility in or out of the selected state, by year, from 1975-2010.

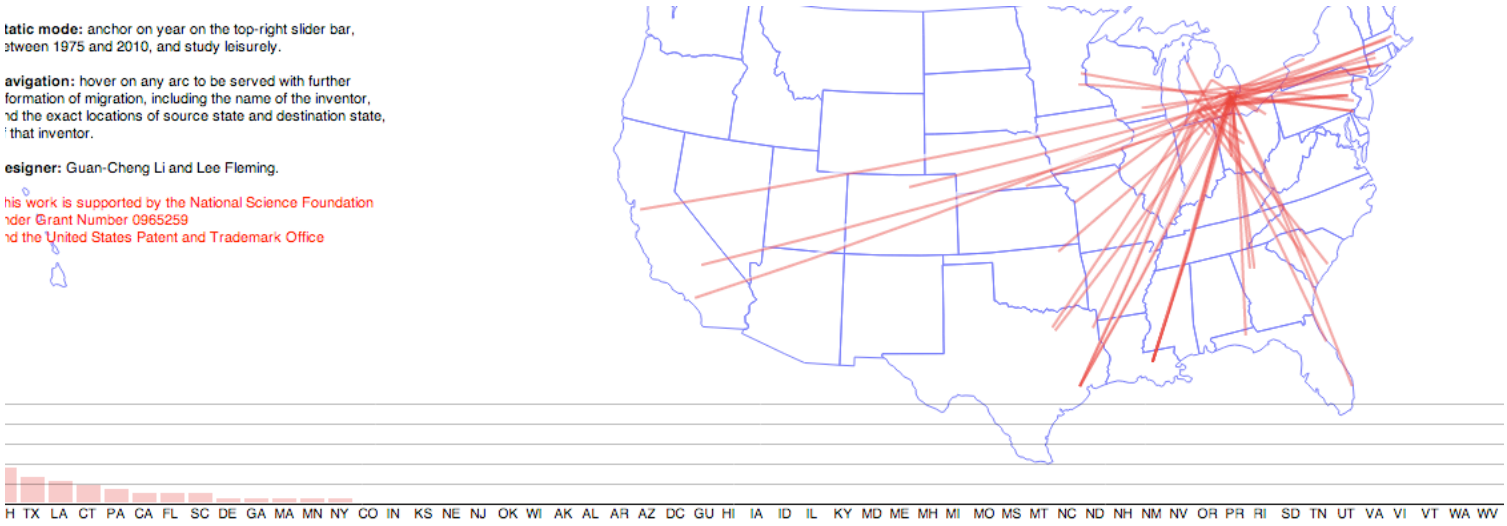

Figure 1: Emigration of patented inventors from Michigan in 1982.

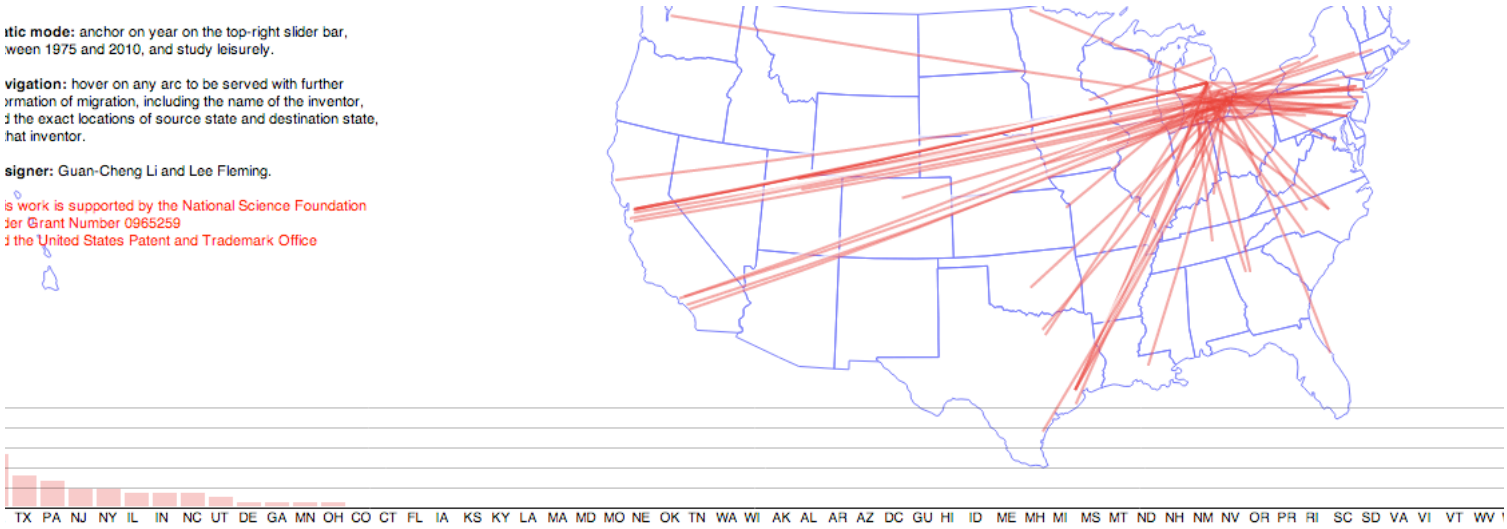

Figure 2: Emigration of patented inventors from Michigan in 1987. Note the greater total amount of emigration (the right hand tail of the distribution represents one inventor in both cases), along with the greater proportion to California, Washington, and Minnesota, states that do not enforce noncompete covenants. For comprehensive evidence, please see Marx et al. 2012.

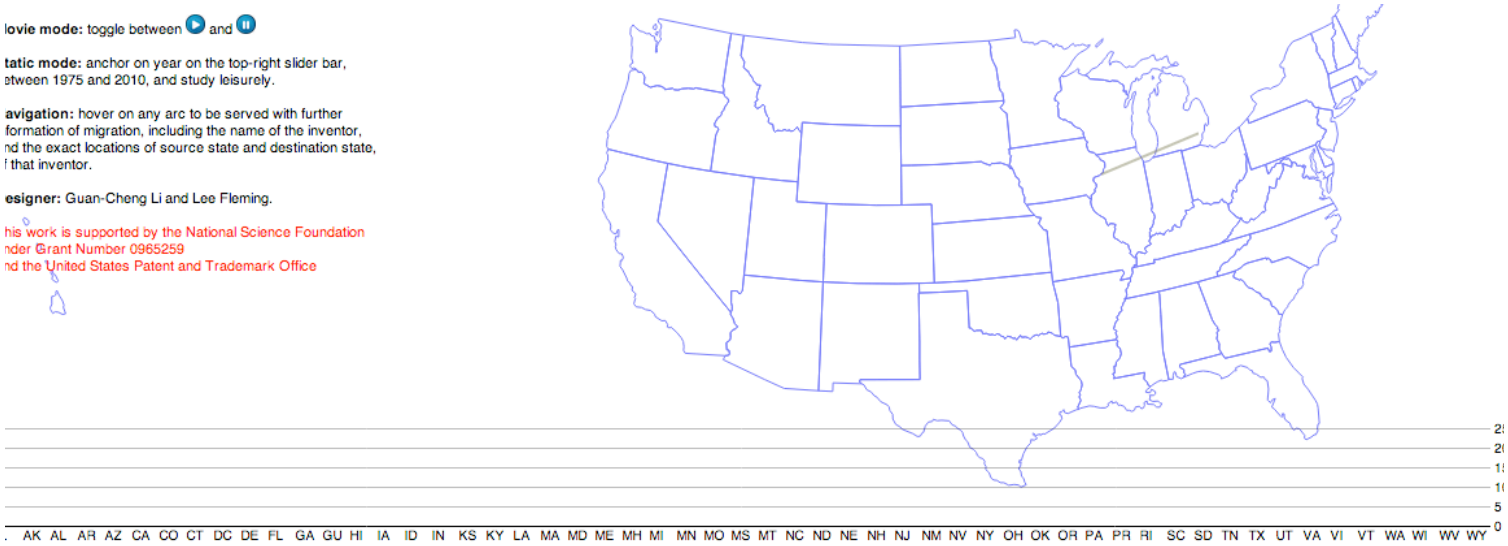

Figure 3: Immigration of patented inventors into Michigan in 1987 (one inventor moved to Michigan, from Illinois). Note the stark contrast with emigration (Figure 2).

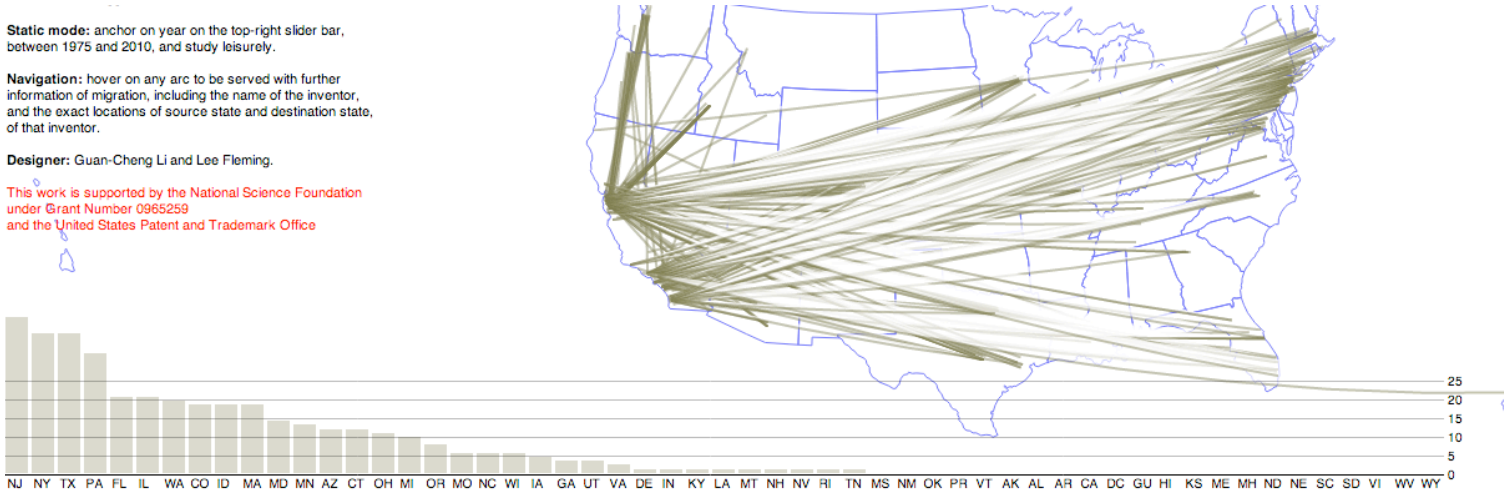

Figure 4: Immigration of patented inventors into California in 2000, during a technology boom.

Implementation of the Interactive Mobility Map

The web-based visualization is coded in JavaScript as well as SVG (Scalable Vector Graphics), and makes use of assisting libraries including D3.js (Bostock et al. 2011), jQuery, and GeoJSON (Geo JavaScript Object Notation). The data exchange from the source of patents in a spreadsheet to JSON is detailed as follows.

We downloaded the disambiguated patents from Patent Network Dataverse (IQSS Harvard): http://dvn.iq.harvard.edu/dvn/dv/patent and retrieved invpat.csv, which contains 9,358,183 patent-inventor instances registered at USPTO between 1975 and 2010 (a patent-inventor instance occurs for each inventor on each patent – for example, a sole inventor creates one patent-inventor instance, and a patent with three authors creates three instances, see Lai et al. 2010). Table 1 provides an example of the invpat.csv file that contains the following fields:

Firstname	Lastname	City	State	Country	Zipcode	InvSeq	Patent	GYear	AppYearStr
PHILIP E	DURAND	HUDSON	MA	US	1749	1	3858241	1975	1974
LONNIE H	NORRIS	MILFORD	MA	US	1757	2	3858241	1975	1974
ELWYN R	GOODING	PINCKNEY	MI	US	48169	1	3858242	1975	1973
RICHARD L	MANN	WOODSTO(	CT	US	6281	1	3858244	1975	1973
MICHAEL A	NATE II	NEW YORK	NY	US	10292	1	3858245	1975	1972
MAURICE A	MANN	NEW YORK	NY	US	10292	2	3858245	1975	1972
SIMCHA	MILO	AUSTIN	TX	US	78799	1	3858246	1975	1973

AppDateStr	Assignee	AsgNum	Class	Invnum	lower	upper
03/26/1974	UNITED STATES	H000000000	2	03858241-1	03858241-	03858241-1
03/26/1974	UNITED STATES	H000000000	2	03858241-2	03858241-	03858241-2
04/16/1973			2	03858242-1	03858242-	03858242-1
12/17/1973			2/69	03858244-1	03858244-	03858244-1
05/18/1972	HAIR AGAIN LTD	H000000108	128/606/6	03858245-1	03858245-	03858245-1
05/18/1972	HAIR AGAIN LTD	H000000108	128/606/6	03858245-2	03858245-	03858245-2
04/19/1973			623/137	03858246-1	03858246-	03858246-1

Table 1: Example from invpat.csv.

When calculating mobility of inventors, we scan through the "Country" column, and first remove the rows that have a non-US value. We augment the table by two new columns: latitude and longitude. For each row, we fill in the values of latitude and longitude by that row's zipcode value, by first converting the zipcode to FIPS (Federal Information Processing Standard) code (GEO-ID) using http://www.census.gov/geo/www/gazetteer/files/Gaz_zcta_national.txt .

Then, we convert the FIPS code to latitude and longitude, which locates the center of that FIPS area, by using: http://www.census.gov/geo/www/gazetteer/files/Gaz_counties_national.txt .

This provides a location for every patent. We compute mobility by partitioning the large invpat.csv table down by calendar year, from 1975 to 2010. This requires us to compare the "AppYearStr" value. We iterate from 1975 to 2010, choose a state of interest (in our study, we have somewhat arbitrarily chosen California, Colorado, Illinois, Massachusetts, Maryland, Michigan, Minnesota, New York, New Jersey, Pennsylvania, Texas, Washington), shortlist the rows which have the matched "AppYearStr" and "State", and save this sub-table for use in the next step.

In the sub-table, for each row, we fetch the inventor's name. This requires us to look at values at both "Firstname" and "Lastname". Note that this table has been disambiguated and therefore we don't need to worry about naming confusion. We scan through the original table at invpat.csv. If there is an inventor name match and AppYearStr" match, but the "State" is different, mobility is assumed. Looking further, if the scanned "AppDateStr" value predates the value in the sub-table, we assume that immigration into that state of interest has occurred and establish an in-arc. Conversely, if the scanned "AppDateStr" value postdates the value in the sub-table, we assume emigration has occurred and establish an out arc. Finally, we save the arc information in form of JSON (JavaScript Object Notation) for rendering on the browser, and the visualization is scripted using Javascript and D3.js. The interaction of year-dragging and arc-hovering is coded using jQuery.js.

On the interface, the immigration or emigration of inventors across the country is mapped in sweeping arcs, reminiscent of airline routes, which allows the users to hover and see details (for example, the inventor and the job change that causes the arc can be identified, as illustrated in Figure 5). Apart from the arcs that signify mobility, the users can also toggle between inflow/outflow, switch amongst U.S. states, and drag along the years for exploratory research.

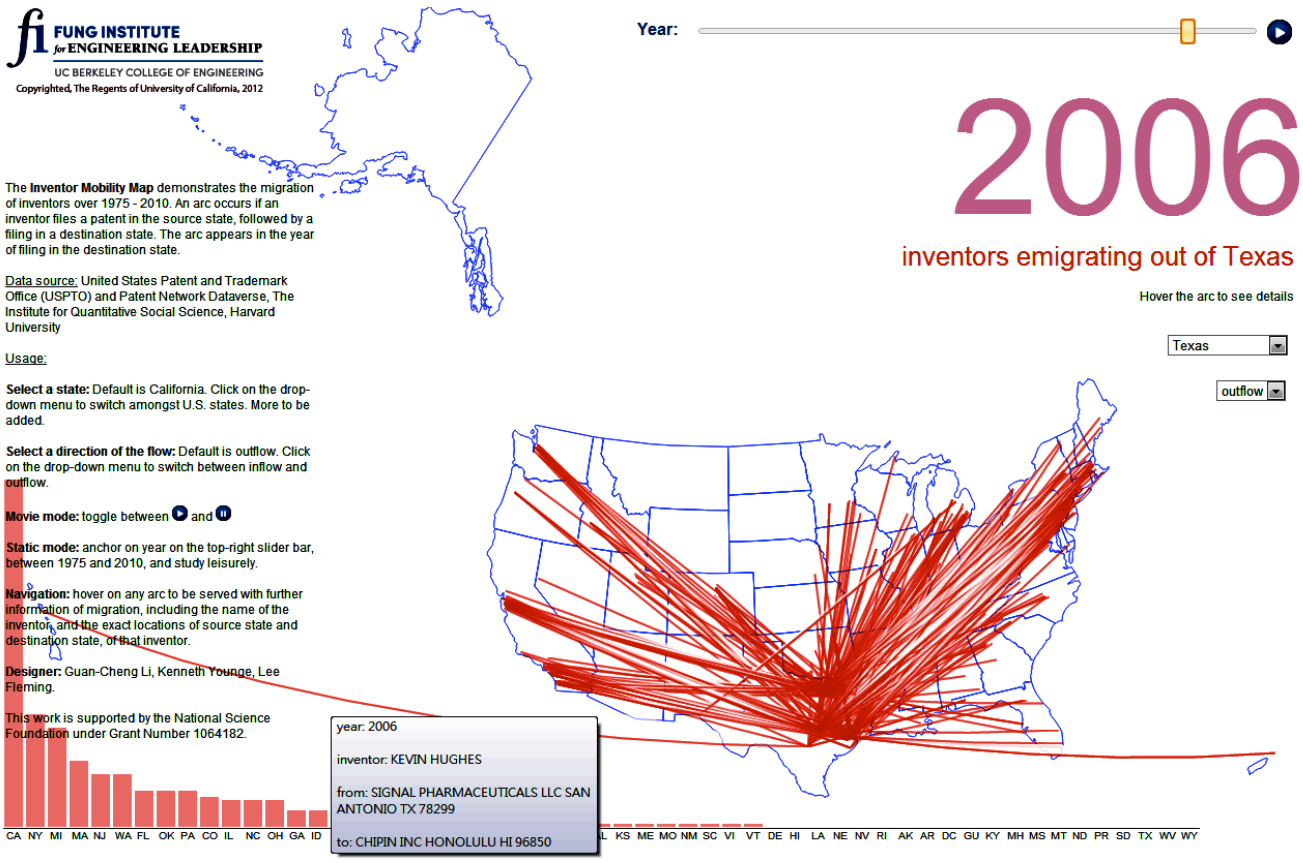

Figure 5: Emigration of patented inventors out of Texas in 2006. As can be seen in the bottom histogram, most went to California, then New York, and then perhaps surprisingly, given the 1987 Michigan graph, Michigan.

References

Agrawal, A. and I. Cockburn and J. McHale. 2006. Gone But Not Forgotten: Labor Flows, Knowledge Spillovers and Enduring Social Capital; *Journal of Economic Geography*; Issue: 6 (5); Pages: 571-591.

Almeida, P. and B. Kogut. 1999. "The Mobility of Engineers in Regional Networks." *Management Science*, Vol. 45, No. 7, July 1999. 907-917.

Bostock, M., Ogievetsky, V., and Heer, J. D3: Datadriven documents, *IEEE Transaction on Visualization and Computer Graphics (Proc. InfoVis)*, 2011.

Breschi, S. and F. Lissoni. 2009. "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows." *Journal of Economic Geography*;Jul2009, Vol. 9 Issue 4, p439.

Lai, R. Alexander D'Amour; Amy Yu; Ye Sun; Lee Fleming, 2010, "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)". Working paper, Fung Institute: http://funginstitute.berkeley.edu/sites/default/files/Disambiguation%20and%20Co-authorship%20Networks%20of%20the%20U.S.%20Patent%20Inventor%20Database%20%281975-2010%29.pdf

Marx, M. and D. Strumsky, L. Fleming "Mobility, Skills, and the Michigan Non-compete Experiment," *Management Science*, 55 (2009): 875-889.

Marx, M. and J. Singh, 2012. "Regional Disadvantage: Non-competes and Brain Drain." Working paper, Fung Institute: http://funginstitute.berkeley.edu/sites/default/files/Regional%20Disadvantage%3F%20Non-Compete%20Agreements%20and%20Brain%20Drain.pdf