NBER WORKING PAPER SERIES

THE AGGLOMERATION OF US ETHNIC INVENTORS

William Kerr

Working Paper 15501
http://www.nber.org/papers/w15501

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2009

Comments are appreciated and can be sent to wkerr@hbs.edu. I am grateful to William Lincoln and Debbie Strumsky for data assistance. I thank Ed Glaeser, Jeff Furman, and participants in the NBER Economics of Agglomeration conferences for helpful comments and research ideas. This research is supported by Harvard Business School, the National Science Foundation, the Innovation Policy and the Economy Group, and the MIT George Schultz Fund. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Agglomeration of US Ethnic Inventors
William Kerr
NBER Working Paper No. 15501
November 2009
JEL No. F15,F22,J44,J61,O31

ABSTRACT

The ethnic composition of US inventors is undergoing a significant transformation, with deep impacts for the overall agglomeration of US innovation. This study applies an ethnic-name database to individual US patent records to explore these trends with greater detail. The contributions of Chinese and Indian scientists and engineers to US technology formation increase dramatically in the 1990s. At the same time, these ethnic inventors became more spatially concentrated across US cities. The combination of these two factors helps stop and reverse long-term declines in overall inventor agglomeration evident in the 1970s and 1980s. The heightened ethnic agglomeration is particularly evident in industry patents for high-tech sectors, and similar trends are not found in institutions constrained from agglomerating (e.g., universities, government).

William Kerr
Harvard Business School
Rock Center 212
Soldiers Field
Boston, MA 02163
and NBER
wkerr@hbs.edu

1 Introduction

Economists have long been interested in agglomeration and innovation. In his seminal outline of the core rationales for industrial clusters, Marshall (1920) emphasized the theory of intellectual spillovers by arguing that in agglomerations, "the mysteries of the trade become no mystery, but are, as it were, in the air." Workers can learn skills quickly from each other in an industrial cluster, and this proximity can speed the adoption of new technologies or best practices. Glaeser and Kahn (2001) argue that the urbanization of high human-capital industries, like finance, is evidence for the role that density plays in the transfer of ideas, and studies of patent citations highlight the importance of local proximity for scientific exchanges (e.g., Jaffe et al. 1992, Thompson and Fox-Kean 2006). Moreover, evidence suggests that agglomeration increases the rate of innovation itself. Saxenian (1994) describes how entrepreneurial firms locate near one another in Silicon Valley to foster new technology development. Carlino et al. (2006) show that higher urban employment density is correlated with greater patenting per capita within cities.

Strong quantitative assessments of the magnitudes and characteristics of intellectual spillovers and agglomeration are essential. Such studies inform business managers of the advantages and costs for locating in areas that are rich in ideas but most likely come with higher rents and wages as well. Moreover, these studies are important for understanding short-run and long-run urban growth and development. They help inform whether industrial specialization or diversity better foster regional development (e.g., Jacobs 1970, Glaeser et al. 1992, Henderson et al. 1995, Duranton and Puga 2001, Duranton 2007) and the role of local knowledge development and externalities in generating sustained growth (e.g., Romer 1986, 1990, Furman et al. 2002). Rosenthal and Strange (2003) note that intellectual spillovers are strongest at the very local levels of proximity.[1]

This study contributes to our empirical understanding of agglomeration and innovation by documenting patterns in the city-level agglomeration of ethnic inventors (e.g., Chinese, Indian) within the US from 1975 through 2007. The contributions of these immigrant groups to US technology formation are staggering: while foreign-born account for just over 10% of the US working population, they represent 25% of the US science and engineering (SE) workforce and nearly 50% of those with doctorates. Even looking within the Ph.D. level, ethnic researchers make exceptional contributions to science as measured by Nobel Prizes, elections to the National Academy of Sciences, patent citation counts, and so on.[2] Recent work relates immigration and growth in US invention (e.g., Peri 2007, Hunt 2008, Kerr and Lincoln 2008). Moreover, ethnic

[1]Several studies assess the relative importance of intellectual spillovers versus other rationales for industrial agglomeration (e.g., lower transportation costs, labor market pooling). Representative papers include Audretsch and Feldman (1996), Rosenthal and Strange (2001), Henderson (2003), Ellison et al. (2007), and Glaeser and Kerr (2008). Porter (1990) emphasizes how vertically related industries may co-locate for knowledge sharing.

[2]For example, Stephan and Levin (2001), Burton and Wang (1999), Johnson (1998, 2001), and Streeter (1997).

entrepreneurs are very active in commercializing new technologies, especially in high-tech sectors (e.g., Saxenian 2002a, Wadhwa et al. 2007).

The spatial distribution of ethnic inventors across US cities, however, is not uniform or random. This agglomeration reflects the general tendency of both high-skilled and low-skilled immigrants to concentrate in certain US cities. Larger cities are often favored for their greater opportunities for assimilation. Geographical distances of cities to home countries and past immigration networks are also important for location decisions. Edin et al. (2003) and Pedace and Rohn (2008) provide recent evidence on the employment effects of enclaves at both the city and sub-city levels. A number of studies in labor economics use spatial differences across cities and occupations in immigrant shares to estimate the impact of higher immigration rates on native workers (e.g., Card 1990, 2001).[3]

The study of how US ethnic inventors agglomerate is thus very important given 1) the disproportionate contributions of immigrant researchers and 2) their non-random spatial distribution across the US. Such a characterization is necessary for understanding the geography of US innovation and economic growth. Moreover, the spatial variation of immigrant researchers across cities allows for stronger quantitative assessments of the role of innovation in city growth. This paper is a first step in this direction.

Econometric studies quantifying the role of ethnic scientists and engineers for technology formation and diffusion are often hampered, however, by data constraints. It is very difficult to assemble sufficient cross-sectional and longitudinal variation for large-scale panel exercises.[4] This paper describes a new approach for quantifying the ethnic composition of US inventors with previously unavailable detail. The technique exploits the inventor names contained on the micro-records for all patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to May 2008.[5] Each patent record lists one or more inventors, with 8 million inventor names associated with the 4.5 million patents. The USPTO grants patents to inventors living within and outside of the US, with each group accounting for about half of patents over the 1975-2008 period.

This study maps into these inventor names an ethnic-name database typically used for commercial applications. This approach exploits the idea that inventors with the surnames Chang or Wang are likely of Chinese ethnicity, those with surnames Rodriguez or Martinez of Hispanic

[3] General surveys of immigration include Borjas (1994), Friedberg and Hunt (1995), Freeman (2006), and Kerr and Kerr (2008).

[4] While the decennial Census provides detailed cross-sectional descriptions, its longitudinal variation is necessarily limited. The annual Current Population Survey, however, provides poor cross-sectional detail and does not ask immigrant status until 1994. The SESTAT database offers a better trade-off between the two dimensions, but suffers important sampling biases with respect to immigrants (Kannankutty and Wilkinson 1999).

[5] The project initially employed the NBER Patent Data File, compiled by Hall et al. (2001), that includes patents granted by the USPTO from January 1975 to December 1999. The current version now employs an extended version developed by HBS Research that includes patents granted through May 2008.

ethnicity, and so on. The match rates are 92%-98% for US domestic inventor records, depending upon the procedure employed, and the process affords the distinction of nine ethnicities: Chinese, English, European, Hispanic/Filipino, Indian/Hindi, Japanese, Korean, Russian, and Vietnamese. Moreover, because the matching is done at the micro-level, greater detail on the ethnic composition of inventors is available annually on multiple dimensions: technologies, cities, companies, and so on. Section 2 describes this data development in greater detail.

Section 3 then documents the growing contribution of ethnic inventors to US technology formation. The rapid increase during the 1990s in the share of high-tech patents granted to Chinese and Indian inventors is particularly striking. This section also uses the patenting data to calculate concentration indices for US innovation. Ethnic inventors have higher levels of spatial concentration than English inventors throughout the thirty-year period studied. Moreover, the spatial concentration of ethnic inventors increases significantly from 1995 to 2004, especially in high-tech sectors like computer-related patenting. The combination of greater ethnic shares and increasing agglomeration of ethnic inventors helps stop and reverse the 1975-1994 declines in the overall concentration of US invention. These trends are confined to industrial patents; universities and government bodies — that are constrained from agglomerating — do not show recent increases in spatial clustering.

The final section concludes. The higher agglomeration of immigrants in cities and occupations has long been noted. For example, Mandorff (2007) highlights how immigrant entrepreneurs tend to agglomerate in selected industries, a process that increases their business impact for specific sectors. Examples within the US are Korean entrepreneurs in dry cleaning, Vietnamese in nail salons, Gujarati Indians in traveler accommodations, Punjabi Indians in gas stations, Greeks in restaurants, and so on. The higher natural social interactions among these ethnic groups aid in the acquisition and transfer of sector-specific skills; scale economies lead to occupational clustering by minority ethnic groups.

To date, there has been very little work, theoretically or empirically, on the agglomeration of US ethnic scientists and engineers with the notable exception of Agrawal et al. (2007).[6] This scarcity of research is disappointing given the scale of these ethnic contributions and the importance of innovation to regional economic growth. Moreover, the large shifts in ethnic inventor populations, often driven in part by US immigration restrictions, may provide empirical footholds for testing agglomeration theories in a natural experiment framework. It is hoped that the empirical platform developed in this study provides a foothold for furthering such analyses.

[6] Agrawal et al. (2007) jointly examine knowledge diffusion through co-location and co-ethnicity using domestic patent citations made by Indian inventors living in the US. While being in the same city or the same ethnicity both encourage knowledge diffusion, their estimations suggest that the marginal benefit of co-location is four times larger for inventors of different ethnicities. This substitutability between social and geographic proximity can create differences between a social planner's optimal distribution of ethnic members and what the inventors themselves would choose.

2 Ethnic-Name Matching Technique

This section describes the ethnic-name matching strategy, outlines the strengths and weaknesses of the name database selected, and offers some validation exercises using patent records filed by foreign inventors with the USPTO. Kerr (2007) further describes the name-matching process, the international name distribution technique, and the apportionment of non-unique matches that are highlighted below.

2.1 Melissa Ethnic-Name Database and Name-Matching Technique

The ethnic-name database employed in this study was originally developed by the Melissa Data Corporation for use in direct-mail advertisements. Ethnic-name databases suffer from two inherent limitations — not all ethnicities are covered and included ethnicities usually receive unequal treatment. The strength of the Melissa database is in the identification of Asian ethnicities, especially Chinese, Indian/Hindi, Japanese, Korean, Russian, and Vietnamese names. The database is comparatively weaker for looking within continental Europe. For example, Dutch surnames are collected without first names, while the opposite is true for French names. The Asian comparative advantage and overall cost effectiveness led to the selection of the Melissa database, as well as the European amalgamation employed in the matching technique. In total, nine ethnicities are distinguished: Chinese, English, European, Hispanic/Filipino, Indian/Hindi, Japanese, Korean, Russian, and Vietnamese.[7]

The second limitation is that commercial databases vary in the number of names they contain for each ethnicity. These differences reflect both uneven coverage and that some ethnicities are more homogeneous in their naming conventions. For example, the 1975 to 1999 Herfindahl indices of foreign inventor surnames for Korean (0.047) and Vietnamese (0.112) are significantly higher than Japanese (0.013) and English (0.016) due to frequent Korean surnames like Kim (16%) and Park (12%) and Vietnamese surnames like Nguyen (29%) and Tran (12%).

Two polar matching strategies are employed to ensure coverage differences do not overly influence ethnicity assignments.

> *Full Matching:* This procedure utilizes all of the name assignments in the Melissa database and manually codes any unmatched surname or first name associated with 100 or more inventor records. This technique further exploits the international distribution of inventor names within the patent database to provide superior results.

[7] The largest ethnicity in the US SE workforce absent from the ethnic-name database is Iranian, which accounted for 0.7% of bachelor-level SEs in the 1990 Census.

The match rate for this restricted procedure is 98% (98% US, 98% foreign). This rate should be less than 100% with the Melissa database as not all ethnicities are included.

Restricted Matching: A second strategy employs a uniform name database using only the 3000 and 200 most common surnames and first names, respectively, for each ethnicity. These numerical bars are the lowest common denominators across the major ethnicities studied. The match rate for this restricted procedure is 89% (92% US, 86% foreign).

For matching, names in both the patent and ethnic-name databases are capitalized and truncated to ten characters. Approximately 88% of the patent name records have a unique surname, first name, or middle name match in the Full Matching procedure (77% in the Restricted Matching), affording a single ethnicity determination with priority given to surname matches. For inventors residing in the US, representative probabilities are assigned to non-unique matches using the masters-level SE communities in Metropolitan Statistical Areas (MSAs). Ethnic probabilities for the remaining 3% of records (mostly foreign) are calculated as equal shares.

2.2 Inventors Residing in Foreign Countries and Regions

Visual confirmation of the top 1000 surnames and first names in the USPTO records confirms the name-matching technique works well. The appendix documents the 100 most common surnames of US-based inventors for each ethnicity, along with their relative contributions. These counts sum the ethnic contribution from inventors with each surname. These counts include partial or split assignments. Moreover, they are not necessarily direct or exclusive matches (e.g., the ethnic match may have occurred through the first name). While some inventors are certainly misclassified, the measurement error in aggregate trends building from the micro-data is minor. The Full Matching procedure is the preferred technique and underlies the trends presented in the next section, but most applications find negligible differences when the Restricted Matching dataset is employed instead.

The application of the ethnic-name database to the inventors residing outside of the US provides a natural quality-assurance exercise for the technique. Inventions originating outside the US account for just under half of USPTO patents, with applications from Japan comprising about half of this foreign total. The appendix documents the results of applying the ethnic-matching procedures for countries and regions grouped to the ethnicities identifiable with the database. The results are very encouraging. First, the Full Matching procedure assigns ethnicities to a large percentage of foreign records, with the match rates greater than 93% for all countries. In the Restricted Matching procedure, a matching rate of greater than 74% holds for all regions.

Second, the estimated inventor compositions are reasonable. The own-ethnicity shares are summarized in the fourth and fifth columns. The weighted average is 86% in the Full Matching procedure, and own-ethnicity contributions are greater than 80% in the UK, China, India, Japan, Korea, and Russia regardless of the matching procedure employed. Like the US, own-ethnicity contributions should be less than 100% due to foreign researchers. The high success rate using the Restricted Matching procedure indicates that the ethnic-name database performs well without exploiting the international distribution of names, although power is lost with Europe. Likewise, uneven coverage in the Melissa database is not driving the ethnic composition trends.

2.3 Advantages and Disadvantages of Name-Matching Technique

The matched records describe the ethnic composition of US scientists and engineers with previously unavailable detail: incorporating the major ethnicities working in the US SE community; separating out detailed technologies and manufacturing industries; providing city-level statistics; and providing annual metrics. Moreover, the assignment of patents to corporations and institutions affords firm-level and university-level characterizations that are not otherwise possible (e.g., the ethnic composition of IBM's inventors filing computer patents from San Francisco in 1985). The next section studies the agglomeration of invention along these various dimensions.[8]

The ethnic-name procedure does, however, have two potential limitations for empirical work on agglomeration that should be highlighted. First, the approach does not distinguish foreign-born ethnic researchers in the US from later generations working as SEs. The procedure can only estimate total ethnic SE populations, and concentration levels are to some extent measured with time-invariant error due to the name-matching approach. The resulting data are very powerful, however, for panel econometrics that employ changes in these ethnic SE populations for identification. Moreover, Census and INS records confirm Asian changes are primarily due to new SE immigration for this period, substantially weakening this concern when examining these groups.

The name-matching technique also does not distinguish finer divisions within the nine major ethnic groupings. For some analyses (e.g., network ties), it would be advantageous to separate Mexican from Chilean scientists within the Hispanic ethnicity, to distinguish Chinese engineers with ethnic ties to Taipei versus Beijing versus Shanghai, and so on. These distinctions are not possible with the Melissa database, and researchers should understand that measurement error

[8] Sample applications are Kerr (2008a,b), Kerr and Lincoln (2008), and Foley and Kerr (2008).

from the broader ethnic divisions may bias their estimated coefficients downward depending upon the application. Nevertheless, the upcoming sections demonstrate how the deep variation available with the ethnic patenting data provides a rich description of US ethnic invention.

3 The Agglomeration of US Ethnic Invention

This section starts by describing the broad trends in ethnic contributions to US technology formation. The spatial concentration of ethnic invention is then closely analyzed, including variations by technology categories and institutions.

3.1 Ethnic Composition of US Inventors

Table 1 describes the ethnic composition of US inventors for 1975-2004, with granted patents grouped by application years. The trends demonstrate a growing ethnic contribution to US technology development, especially among Chinese and Indian scientists. Ethnic inventors are more concentrated in high-tech industries like computers and pharmaceuticals and in gateway cities relatively closer to their home countries (e.g., Chinese in San Francisco, Europeans in New York, and Hispanics in Miami). The final three rows demonstrate a close correspondence of the estimated ethnic composition to the country-of-birth composition of the US SE workforce in the 1990 Census. The estimated European contribution in Table 1 is naturally higher than the immigrant contribution measured by foreign born.

Figure 1 illustrates the evolving ethnic composition of US inventors from 1975-2004. The omitted English share declines from 83% to 70% during this period. Looking across all technology categories, the European ethnicity is initially the largest foreign contributor to US technology development. Like the English ethnicity, however, the European share of US domestic inventors declines steadily from 8% in 1975 to 6% in 2004. This declining share is partly due to the exceptional growth over the thirty years of the Chinese and Indian ethnicities, which increase from under 2% to 8% and 5%, respectively. As shown below, this Chinese and Indian growth is concentrated in high-tech sectors, where Chinese inventors supplant European researchers as the largest ethnic contributor to US technology formation. The Indian ethnic contribution declines somewhat after 2000.[9]

Among the other ethnicities, the Hispanic contribution grows from 3% to 4% from 1975 to 2004. The level of this series is likely mismeasured due to the extensive overlap of Hispanic and

[9] This decline is mostly due to changes within the computer technology sector as seen below. Recent applications to the USPTO suggest the Indian trend may not have declined as much as the granted patents through early 2008 portray. Kerr and Lincoln (2008) investigate the role of H-1B visa reforms for explaining these patterns.

European names, but the positive growth is consistent with stronger Latino and Filipino scientific contributions in Florida, Texas, and California. The Korean share increases dramatically from 0.3% to 1.1% over the thirty years, while the Russian climbs from 1.2% to 2.2%. Although difficult to see with Figure 1's scaling, much of the Russian increase occurs in the 1990s following the dissolution of the Soviet Union. The Japanese share steadily increases from 0.6% to 1.0%. Finally, while the Vietnamese contribution is the lowest throughout the sample, it does exhibit the strongest relative growth from 0.1% to 0.6%.

The 1975-2004 statistics employ patents granted by the USPTO through May 2008. Due to the long and uneven USPTO review process, statistics are grouped by application year to construct the most accurate indicators of when inventive activity occurs. The unfortunate consequence of using application years, however, is substantial attrition in years immediately before 2008. As many patents are in the review process but have yet to be granted, the granted patent series is truncated at the 2004 application year. The USPTO began publishing patent applications in 2001. These applications data also show comparable ethnic contributions.

3.2 Spatial Locations of US Ethnic Inventors

Table 2 examines the 1975-2004 ethnic inventor contributions by major MSAs. A total of 283 MSAs are identified from inventors' city names using city lists collected from the Office of Social and Economic Data Analysis at the University of Missouri, with a matching rate of 99%. Manual coding further ensures all patents with more than 100 citations and all city names with more than 100 patents are identified. The first four columns document each MSA's share of US patenting. Not surprisingly, these shares are highly correlated with MSA size, with the three largest patenting centers for 1995-2004 found in San Francisco (12%), New York (7%), and Los Angeles (6%), where the percentages indicate US domestic patent shares.

Comparing these total patenting percentages with the ethnic patenting shares, listed in the second set of four columns, reveals the more interesting fact that ethnic patenting is more concentrated than general innovation. The 1995-2004 ethnic patent shares of San Francisco, New York, and Los Angeles are 19%, 10%, and 8%, respectively. Similarly, 81% of ethnic research occurs in the major MSAs listed in Table 2, compared to 73% of total patenting. The final three columns list the Chinese and Indian patenting share by MSA, highlighting the exceptional growth of San Francisco from 8% of 1975-1984 patenting to 25% in 1995-2004. These concentration levels and trends are further examined below.[10]

[10] Each of these trends appears to have strengthened in the recent applications data (i.e., the columns marked with A in Table 2). While suggestive, these statistics should be treated with caution. Some technology fields and firm types are more likely to publish their patent applications than others. Likewise, probabilities of patent grants conditional on application vary by field. Lemley and Sampat (2007) discuss these limitations further.

Table 3 presents simple least squares estimations of ethnic inventor locations and MSA characteristics. The variables of interest are MSA shares of US ethnic inventors during 1985-2004, with column headers indicating ethnicities. These shares are calculated over the 244 MSAs for which full covariate information are assembled. The dropped observations are small cities not separately identified in 1990 Census of Population. For ease of interpretation, variables are transformed to have unit standard deviation in these cross-sectional estimations. Estimations are weighted by MSA populations.

To establish a baseline, the first two columns consider MSA inventor shares of the English ethnicity. In Column 1, MSA size and urban density strongly predict higher English inventor shares. A one standard-deviation increase in the population share of the MSA correlates with a 0.57 standard-deviation increase in the share of English ethnic invention. Coastal access does not predict greater inventor concentration in multi-variate frameworks, although a univariate correlation exists. On the other hand, MSA demographics have a statistically and economically significant relationship with inventor concentrations. The MSA traits are calculated from the 1990 Census of Population. MSAs with more-educated workforces are associated with greater inventor concentrations. Higher shares of English invention are also found in MSAs with relatively more people between the ages of 30 and 60 (the omitted group) and more men. All told, this parsimonious set of covariates explains 84% of the variation in English invention shares.

Table 2 suggests that inventor shares are relatively persistent over time for MSAs. Column 2 of Table 3 confirms this observation for English inventors. The estimation incorporates the share of English ethnic patenting in the MSA for 1975-1984. This ten-year period pre-dates the major growth in ethnic inventors highlighted in Figure 1. The spatial distribution of English invention over 1975-1984 is a very strong predictor for 1985-2004 concentration with an elasticity of 0.84. MSA populations and density levels do not exhibit a well-measured relationship with 1985-2004 English inventor concentrations after controlling for these past levels. Partial correlations with MSA demographics, however, are more robust. Incorporating the past concentration lag explains 88% of the MSA-level variation in inventor shares (83% by itself).[11]

The subsequent eight columns of Table 3 consider major non-English inventor shares. The estimation framework remains the same excepting the 1975-1984 MSA inventor shares in the even-numbered columns that are adjusted to match the dependent variable. Most explanatory variables (e.g., MSA demographics) demonstrate similar elasticities across ethnic groups. Coastal access tends to be more important, although of borderline statistical significance. This reflects the well-known tendency for immigrants to locate in port cities closer to their home countries.

[11] Unreported specifications further incorporate mean wages in manufacturing, mean family income levels, and mean housing prices by MSA. Positive correlations between inventor shares and manufacturing wages are generally found; family income levels and housing prices do not exhibit robust relationships in multi-variate settings. The inclusion of these three covariates has very limited influence on the reported outcomes.

Several interesting differences, however, emerge. First, the overall explanatory power of these regressors varies across ethnic groups. The R^2 values for the Chinese and Indian ethnicities are substantially lower than those for the European and Hispanic ethnicities. These Asian ethnicities thus have more idiosyncratic spatial patterns than this limited set of covariates modelled. This is confirmed when the even-numbered columns incorporate the lagged ethnic inventor shares. The gain in the variation explained through past MSA-specific placements is strongest for Chinese and Indian inventors. This strength suggests that lagged spatial patterns for Asian inventors may offer an empirical foothold for predicting future MSA-level innovation even conditional on other MSA-level traits.

These even-numbered columns also show that lagged ethnic inventor shares tend to have weaker predictive power for subsequent MSA-level concentration compared to the English ethnicity in Column 2. The elasticities range from 0.87 for Chinese patents to 0.53 for Hispanic patents (which is lowest among the nine ethnic inventor groups). This lower explanatory power has at least two explanations. First, spatial distributions for ethnic inventors over 1975-1984 may have greater measurement error than English inventor distributions due to smaller counts of relevant patents. Such measurement error would downward bias estimated elasticities.

Nonetheless, it is also true that ethnic inventors facilitate shifts in invention locations across US MSAs. For example, immigrant SE students graduating from elite US universities enter a national labor market. Hispanic inventors have supported broader growth in Florida and the southwestern states. While past immigration cities are favored, ethnic inventors also have an inherent capacity to facilitate regional adjustments. Unreported estimations further test this conclusion by controlling simultaneously for each MSA's 1975-1984 English inventor share and ethnic-specific inventor share. With the exception of the European and Russian ethnicities, lagged ethnic spatial distributions have stronger predictive power for subsequent agglomeration than lagged English spatial distributions.

Table 4 repeats the estimations without the MSA population weights. The measured partial correlations decline in magnitude somewhat, reflective of the greater attention paid to smaller MSA shares, but the patterns of coefficients and explanatory power are comparable to the weighted outcomes. Several additional specification checks are also undertaken. Incorporating regional fixed effects finds anticipated spatial patterns — Midwestern US MSAs tend to have higher invention rates conditional on the covariates modelled, while southern MSAs have lower rates. The east and west coasts are often not statistically distinguishable from each other conditionally. Performing the share estimations on an annual basis, which circumvents growth in recent patent application rates, yields similar outcomes to the cross-sectional results. Likewise, log specifications produce outcomes similar to the share specification framework.

Finally, the appendix documents specifications that model lagged ethnic population shares across MSAs as the historical regressor rather than the distribution of lagged ethnic patenting. These shares are calculated over working-age populations for 203 cities through the 1980 Census of Population by country of birth. In general, the spatial distribution of lagged ethnic patenting in Tables 3 and 4 is a stronger predictor than general ethnic population distributions; R^2 values also decline. The one exception is for the Chinese ethnicity, where the general Chinese population distribution is an exceptionally strong predictor of recent patenting. These patterns also hold when jointly modelling the lagged regressors together.

These comparisons are interesting in that they begin to quantify the relative roles of production versus consumption benefits for the agglomeration of ethnic inventors. The productive benefits of being near other inventors of one's ethnicity appear stronger that the general consumption benefits of being in ethnic enclaves, but the latter are surprisingly strong. To address properly this issue, future work hopes to examine the sub-city level to the extent possible with the patenting data. The high correlation between lagged Chinese inventor and population distributions depends, for example, on the decision to model the San Francisco Bay area as a single MSA. Splitting San Jose and Silicon Valley from San Francisco and/or Oakland would reduce the correlation. Undertaking such an analysis would be informative for the specific question of location decisions by ethnic inventors; it would also contribute to recent work on ethnic enclaves at the sub-city level (e.g., Pedace and Rohn 2008).

Of course, these estimations must be interpreted as partial correlations rather than causal assessments. Clearly, ethnic inventors directly influence many of the determinants modelled (e.g., education shares) and may also have local spillover effects through their work (e.g., local technology gains that generate city population growth). Omitted factors may also be correlated with past immigrant placements. Future work hopes to further refine these determinants in a causal assessment.

Ongoing research is further evaluating how shifts in the geographic concentration of ethnic inventors facilitate changes in the geographic composition of US innovation. Not only are ethnic scientists disproportionately concentrated in major MSAs, but growth in a MSA's share of ethnic patenting is highly correlated with growth in its share of total US patenting. Annual regressions across the full 1975-2004 MSA sample find that an increase of 1% in an MSA's ethnic patenting share correlates with a 0.6% increase in the MSA's total invention share. This coefficient is remarkably high, as the mean ethnic share of total invention during this period is around 20%. Of course, additional study is required before causal assessments are possible. The ethnic-name approach will also need to be complemented with external data to distinguish ethnic inventor shifts due to new immigration, domestic migration, or occupational changes.

3.3 Spatial Concentration of US Ethnic Inventors

To refine the earlier visual observations made regarding agglomeration levels in Table 2, Table 5 presents three concentration indices for US domestic patenting. The first concentration metric studied is the Herfindahl-Hirschman Index defined by $HHI_t = \sum_{m=1}^{M} Share_{mt}^2$, where M indexes 283 MSAs and $Share_{mt}$ is the share of patenting in MSA m in period t. Of course, patenting is undertaken outside of MSAs, too. The share of patenting outside of these 283 MSAs declines from 9% in 1975-1984 to 7% in 1995-2004. In 2001-2006 applications, this share further declines to 6%. This portion of US invention is excluded from the remainder of this paper, with concentration metrics being calculated over MSA patenting only.

The top panel of Table 5 and Figure 2 highlight several important levels differences. First, US invention is more concentrated than the general population across these MSAs.[12] Moreover, ethnic inventors are substantially more agglomerated than English-ethnicity inventors throughout the thirty years considered. The mean population HHI is 0.024 over the period, compared with 0.037 for invention and 0.059 for all non-English inventors. The agglomeration of Chinese inventors further stands out at 0.081. This higher ethnic concentration certainly reflects the well-known concentration of immigrant groups, but is not due to simply the smaller sizes of some ethnicities. Chinese, Japanese, and Vietnamese are consistently the most agglomerated of ethnic inventor groups. European and Hispanic inventors are the least concentrated, but all ethnic groups are more agglomerated than the English ethnicity.[13]

Moving from the levels to the trends evident in Table 5 and Figure 2, the HHI for all US inventors consistently declines from 1975-1979 to 1990-1994. This trend is reversed, however, with greater levels of invention agglomeration in 1995-1999 and 2000-2004. This reversal towards greater patenting concentration is not reflected in the overall population shares. Ethnic inventors, however, show a sharp increase in these latter ten years. This upturn is strongest among Asian ethnic groups, with European and Hispanic inventors showing limited change in agglomeration.

A second agglomeration metric is calculated as the share of total US patenting in the Top 5 MSAs for 1975-1984: New York City (12%), Los Angeles (7%), Chicago (6%), Philadelphia (5%), and San Francisco (5%). Boston (4%) and Detroit (3%) have the next two largest shares in 1975-1984. These five MSAs account for about 37% of MSA patenting during this initial period and 34% of total US patenting that includes rural areas. The share accounted for by

[12] MSA populations are calculated through county populations collected in 1977, 1982, 1987, 1992, and 1997. These are mid-points of the five-year increments studied. The 2000-2004 period uses the 1997 MSA population.

[13] Calculations from the 1990 and 2000 Census of Populations find that the aggregate concentration of immigrant SEs is slightly less than the agglomeration of all immigrants. Substantial differences in immigrant shares are evident in larger cities. New York City, Los Angeles, and Miami have larger overall immigration pools relative to SE, while San Francisco, Washington, Boston, and Seattle have greater SE shares.

these five MSAs behaves similarly to the HHI metric, declining until 1990-1994 before growing during 1995-2004. While less formal, this second technique highlights how ethnic agglomeration shifts across the major US MSAs. By 1995-2004, San Francisco (12%) leads New York City (7%) and Los Angeles (6%). Boston and Chicago would complete a new Top 5 MSAs list for 1995-2004.

Our final agglomeration metric is taken from Ellison and Glaeser (1997),

$$\gamma_e^{Agg} = \frac{\sum_{m=1}^{M}(s_{m,e} - x_m)^2}{1 - \sum_{m=1}^{M} x_m^2},$$

where M indexes MSAs. $s_{1,e}, s_{2,e}, \ldots, s_{M,e}$ are the shares of ethnicity e's patenting contained in each of these geographic areas. $x_1, x_2, \ldots, x_M$ are each area's share of population.[14] This metric estimates the agglomeration of invention relative to the baseline established by the MSA populations. If invention is randomly distributed among the population, the Ellison and Glaeser metric will not show concentration. The bottom panel of Table 5 and Figure 3 report these indices. When judged relative to the overall population's distribution, the trends in the agglomeration of invention look a little different. The 1975-1994 periods are found to have fairly consistent levels of concentration, with a strong upturn in the 1995-2004 years. This pattern is predicted by the growing deviations with time in the HHI trends in Panel A.

Following Ellison et al. (2007), the pairwise coagglomeration of invention between ethnicity e_1 and e_2 is analyzed with the simple formula

$$\gamma_{e_1,e_2}^{Coagg} = \frac{\sum_{m=1}^{M}(s_{m,e_1} - x_m)(s_{m,e_2} - x_m)}{1 - \sum_{m=1}^{M} x_m^2}.$$

This index measures the covariance of ethnic invention across MSAs, with the denominator rescaling the covariance to eliminate a sensitivity to the fineness of the geographic breakdown. The coagglomeration indices are contained in the appendix. Coagglomeration among non-English ethnic inventors is substantially higher than between English inventors and these groups. This is especially true among the Asian ethnicities. These coagglomeration measures rise in recent years, behaving similarly to the agglomeration measures when relative to the total population.

3.4 Technology Concentration of US Ethnic Inventors

Figure 4 documents the total ethnic contribution by the six broad technology groups into which patents are often classified: Chemicals, Computers and Communications, Drugs and Medical,

[14] The full Ellison and Glaeser (1997) formula also controls for the HHI index of plant size. This feature is ignored in this examination of individual inventors. The ethnic patenting data do not easily support continuous estimators like Duranton and Overman (2005), although future research hopes to approximate these metrics too.

Electrical and Electronic, Mechanical, and Miscellaneous/Others. The Miscellaneous group includes patents for agriculture, textiles, furniture, and the like. Growth in ethnic patenting is noticeably stronger in high-tech sectors than in more traditional industries. Figures 5 and 6 provide more detailed glimpses within the Chinese and Indian ethnicities, respectively. These two ethnic groups are clearly important contributors to the stronger growth in ethnic contributions among high-tech sectors, where Chinese inventors supplant European researchers as the largest ethnic contributor to US technology formation.[15]

One possible explanation for Table 5's aggregate gains in concentration is compositional shifts in the volume and nature of granted patents, rather than a shift in underlying innovation per se. There has been a substantial increase in the number of patents granted by the USPTO over the last two decades. While this increase is partly due to population growth and higher levels of US innovation, institutional factors also play an important role.[16] The heightened agglomeration may be driven by greater patenting rates by certain technology groups, reflecting either true changes in the underlying innovation rates or simply a greater propensity to seek patent protection. The latter is especially relevant for the recent rise of software patents (e.g., Graham and Mowery 2004). Microsoft, Oracle, and other software companies are among the US's largest firms today in terms of patent applications, but historically this industry did not seek patent protection.

Table 6 considers the geographic concentration of invention that exists within each of the six broad technology groupings. Panel A presents HHI measures calculated over all patents within each technology. The exceptional rebounds for 1995-2004 are strongest within the Computers and Communications and Electrical and Electronic groupings. Drugs and Medical and Mechanical categories also demonstrate weaker gains, while Chemicals and Miscellaneous show steady trends for less spatial agglomeration throughout the 1975-2004 period.

The dual responses within the Computers and Communications and Electrical and Electronic groupings suggest that the greater agglomeration is more of a high-tech phenomena than software in particular. This conclusion is further confirmed in the appendix. In these estimations, agglomeration is calculated for each sub-category within the six broad technology divisions; there are four to nine sub-categories within each division. In both weighted and unweighted estimations, the concentration metrics at the sub-category level behave similarly to Table 6. This robustness highlights that a few isolated technology categories, either pre-existing or entering with recent USPTO additions, are not solely responsible for the patterns evident.

[15] The USPTO issues patents by technology categories rather than by industries. Combining the work of Johnson (1999), Silverman (1999), and Kerr (2008), concordances can be developed to map the USPTO classification scheme to the three-digit industries in which new inventions are manufactured or used. Scherer (1984) and Keller (2002) further discuss the importance of inter-industry R&D flows.

[16] For example, Griliches (1990), Kortum and Lerner (2000), Kim and Marschke (2004), Hall (2005), Jaffe and Lerner (2005).

Panels B and C report similar indices for English and non-English ethnicity inventors. Some of the sharp concentration gains within the Computers and Communications and Electrical and Electronic groupings can be traced to higher agglomeration of the English inventors. The exceptional growth in concentration among non-English ethnic inventors, however, is even more striking. Figure 7 presents the HHI of Computers and Communications patents for selected ethnic groups. The Chinese HHI reaches just less than 0.200 by 2000-2004, while the Indian concentration also grows to 0.141. Note that this concentration growth occurs during a period of growing patent counts.

Ethnic inventors thus pull up the overall patenting concentration in at least three ways. First, ethnic inventors have higher levels of existing concentration and are becoming a larger share of US patenting (Figure 4). Even if their own concentration holds constant, this should lead to an increase in the agglomeration of US patenting. Second, ethnic inventors are themselves becoming more spatially concentrated in high-tech fields. This force also leads to an increase in overall agglomeration levels. Ethnic inventors are also more concentrated in fields that have experienced greater rates of recent patenting, yielding a mechanical link as well.[17]

3.5 Institutional Concentration of US Ethnic Inventors

Patents are granted to several types of institutions. Industrial firms account for about 70% of patents granted from 1980-1997, while government and university institutions are assigned about 4% of patents. Unassigned patents (e.g., individual inventors) represent about 26% of US invention. Public companies account for 59% of the industry patents during this period. With the exception of unassigned patents, institutions are primarily identified through assignee names on patents.

Figure 8 demonstrates that intriguing differences in ethnic scientific contributions also exist by institution type. Over the 1975-2004 period, ethnic inventors are more concentrated in government and university research labs and in publicly-listed companies than in private companies or as unaffiliated inventors. Part of this levels difference is certainly due to immigration visa sponsorships by larger institutions. Growth in ethnic shares are initially stronger in the government and university labs, but publicly-listed companies appear to close the gap by 2004. The other interesting trend in Figure 8 is for private companies, where the ethnic contribution sharply increases in the 1990s. This rise coincides with the strong growth in ethnic entrepreneurship in high-tech sectors.[18]

[17] These effects appear to continue in the 2001-2006 applications data catalogued in Table 2.

[18] Publicly-listed companies are identified from a 1989 mapping developed by Hall et al. (2001). This company list is not updated for delistings or new public offerings. This approach maintains a constant public grouping for reference, but it also weakens the representativeness of the public and private company groupings at the sample extremes for current companies.

Panels A and B of Table 7 document the evolution of the HHI concentration for industry and university/government patenting, respectively. The column headers again indicate different technology groups. Despite having fairly similar levels of spatial concentration, the differences between institutions in the agglomeration trends for patenting are striking. The concentration of invention within universities and governments has either weakened or remained constant in every technology group. The recent gains in industry concentration, on the other hand, are stronger than the aggregate statistics from Table 6. Whereas the recent growth in industry concentration is strongest for Computers and Communications and Electrical and Electronic, the two technology groups show above-average declines for universities and government bodies.

The bottom two panels of Table 7 show the deeper impact of these institutional differences for non-English invention. Ethnic inventors are again very strong drivers for the recent agglomeration increases in industry patenting within high-tech sectors. On the other hand, ethnic inventors are not becoming more geographically agglomerated within universities and government institutions. This even holds true for Chinese and Indian groups within the Computers and Communications and Electrical and Electronic technology sectors. Figures 9 and 10 summarize these differences. As universities and government bodies are more constrained from agglomerating than industrial firms, these differences provide a nice falsification check on the earlier trends and the role of ethnic inventors.[19]

4 Conclusions

Ethnic scientists and engineers are an important and growing contributor to US technology development. The Chinese and Indian ethnicities, in particular, are now an integral part of US invention in high-tech sectors. The magnitude of these ethnic contributions raises many research and policy questions: debates regarding the appropriate quota for H-1B temporary visas, the possible crowding out of native students from SE fields, the brain drain or brain circulation effect on sending countries, and the future prospects for US technology leadership are just four examples.[20] While the answers to these questions must draw from many fields within and outside of economics, valuable insights can be developed through agglomeration theory and empirical studies.

This paper builds a new empirical platform for these research questions by assigning probable ethnicities for US inventors through the inventor names available with USPTO patent records.

[19] Trends in concentration ratios of unassigned inventors fall in between industry and university/government, behaving more closely like the latter. While there is some recent growth in ethnic inventor concentration within this class, the upturn is much weaker than in industrial firms. Figure 8 also highlights that ethnic inventors are a smaller fraction of unassigned patents, leading to a smaller impact on aggregate statistics.

[20] Representative papers are Lowell (2000), Borjas (2005), Saxenian (2002b), and Freeman (2005), respectively.

The resulting data document with greater detail than previously available the powerful growth in US Chinese and Indian inventors during the 1990s. At the same time, these ethnic inventors became more spatially concentrated across US cities. The combination of these two factors helps stop and reverse long-term declines in overall inventor agglomeration evident in the 1970s and 1980s. The heightened ethnic agglomeration is particularly evident in industry patents for high-tech sectors, and similar trends are not found in institutions constrained from agglomerating (e.g., universities, government).

References

[1] Agrawal, Ajay, Devesh Kapur, and John McHale, "Birds of a Feather – Better Together? Exploring the Optimal Spatial Distribution of Ethnic Inventors", NBER Working Paper 12823 (2007).

[2] Audretsch, David, and Maryann Feldman, "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review* 86:3 (1996), 630-640.

[3] Borjas, George, "Do Foreign Students Crowd Out Native Students from Graduate Programs?", NBER Working Paper 10349 (2004).

[4] Borjas, George, "Economics of Immigration", *Journal of Economic Literature* 32 (1994), 1667-1717.

[5] Burton, Lawrence, and Jack Wang, "How Much Does the U.S. Rely on Immigrant Engineers?", NSF SRS Issue Brief (1999).

[6] Card, David, "The Impact of the Mariel Boatlift on the Miami Labor Market", *Industrial and Labor Relations Review* 43:2 (1990), 245-257.

[7] Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration", *Journal of Labor Economics* 19:1 (2001), 22-64.

[8] Duranton, Gilles, "Urban Evolutions: The Fast, the Slow, and the Still", *American Economic Review* 97:1 (2007), 197-221.

[9] Duranton, Gilles, and Henry Overman, "Testing for Localization Using Micro-Geographic Data", *Review of Economic Studies* 72 (2005), 1077-1106.

[10] Duranton, Gilles, and Diego Puga, "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products", *American Economic Review* 91:5 (2001), 1454-1477.

[11] Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment", *The Quarterly Journal of Economics* 118:1 (2003), 329-357.

[12] Ellison, Glenn, and Edward Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 105:5 (1997), 889-927.

[13] Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", NBER Working Paper 13068 (2007).

[14] Foley, C. Fritz, and William Kerr, "US Ethnic Scientists and Foreign Direct Investment Placement", HBS Working Paper (2008).

[15] Freeman, Richard, "Does Globalization of the Scientific/Engineering Workforce Threaten U.S. Economic Leadership?", NBER Working Paper 11457 (2005).

[16] Freeman, Richard, "People Flows in Globalization", *Journal of Economic Perspectives* 20:2 (2006), 145-170.

[17] Friedberg, Rachel, and Jennifer Hunt, "The Impact of Immigrants on Host Country Wages, Employment and Growth", *Journal of Economic Perspectives* 9:2 (1995), 23-44.

[18] Furman, Jeffrey, Michael Porter, and Scott Stern, "The Determinants of National Innovative Capacity", *Research Policy* 31 (2002), 899-933.

[19] Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer, "Growth in Cities", *Journal of Political Economy* 100:6 (1992), 1126-1152.

[20] Glaeser, Edward, and Matthew Kahn, "Decentralized Employment and the Transformation of the American City", NBER Working Paper 8117 (2001).

[21] Glaeser, Edward, and William Kerr, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" forthcoming in *Journal of Economics and Management Strategy* (2008).

[22] Graham, Stuart, and David Mowery, "Software Patents: Good News or Bad News?", Working Paper (2004).

[23] Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature* 28:4 (1990), 1661-1707.

[24] Hall, Bronwyn, "Exploring the Patent Explosion", *Journal of Technology Transfer* 30 (2005), 35-48.

[25] Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498 (2001).

[26] Henderson, J. Vernon, "Marshall's Scale Economies", *Journal of Urban Economics* 53 (2003), 1–28.

[27] Hunt, Jennifer, "How Much Does Immigration Boost Innovation?", Working Paper (2008).

[28] Jacobs, Jane, *The Economy of Cities* (New York, NY: Vintage Books, 1970).

[29] Jaffe, Adam, and Joshua Lerner, *Innovation and Its Discontents* (Boston, MA: Harvard Business School Press, 2005).

[30] Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, 108:3 (1993), 577-598.

[31] Johnson, Daniel, "150 Years of American Invention: Methodology and a First Geographic Application", Wellesley College Economics Working Paper 99-01 (1999).

[32] Johnson, Jean, "Statistical Profiles of Foreign Doctoral Recipients in Science and Engineering: Plans to Stay in the United States", NSF SRS Report (1998).

[33] Johnson, Jean, "Human Resource Contribution to U.S. Science and Engineering From China", NSF SRS Issue Brief (2001).

[34] Kannankutty, Nirmala, and R. Keith Wilkinson, "SESTAT: A Tool for Studying Scientists and Engineers in the United States", NSF SRS Report (1999).

[35] Keller, Wolfgang, "Trade and the Transmission of Technology", *Journal of Economic Growth* 7 (2002), 5-24.

[36] Kerr, Sari Pekkala, and William Kerr, "Economic Impacts of Immigration: A Survey", HBS Working Paper 09-013 (2008).

[37] Kerr, William, "The Ethnic Composition of US Inventors", HBS Working Paper 08-006 (2007).

[38] Kerr, William, "Ethnic Scientific Communities and International Technology Diffusion", *Review of Economics and Statistics* 90:3 (2008a), 518-537.

[39] Kerr, William, "Heterogeneous Technology Diffusion and Ricardian Trade Patterns", HBS Working Paper (2008b).

[40] Kerr, William, and William Lincoln, "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention", HBS Working Paper 09-005 (2008).

[41] Kim, Jinyoung, and Gerald Marschke, "Accounting for the Recent Surge in U.S. Patenting: Changes in R&D Expenditures, Patent Yields, and the High Tech Sector", *Economics of Innovation and New Technologies* 13:6 (2004), 543-558.

[42] Kortum, Samuel, and Joshua Lerner, "Assessing the Contribution of Venture Capital to Innovation", *RAND Journal of Economics* 31:4 (2000), 674-692.

[43] Lemley, Mark, and Bhaven Sampat, "Is the Patent Office a Rubber Stamp?", Stanford Law School Working Paper 999098 (2007).

[44] Lowell, B. Lindsay, "H1-B Temporary Workers: Estimating the Population", The Center for Comparative Immigration Studies Working Paper 12 (2000).

[45] Mandorff, Martin, "Social Networks, Ethnicity, and Occupation", Working Paper (2007).

[46] Marshall, Alfred, *Principles of Economics* (London, U.K.: MacMillan and Co., 1920).

[47] Pedace, Roberto, and Stephanie Rohn, "A Warm Embrace or the Cold Shoulder: Wage and Employment Outcomes in Ethnic Enclaves", CES Working Paper 08-09 (2008).

[48] Peri, Giovanni, "Higher Education, Innovation and Growth", in Giorgio Brunello, Pietro Garibaldi and Etienne Wasmer (eds.) *Education and Training in Europe* (Oxford: Oxford University Press, 2007).

[49] Porter, Michael, *The Competitive Advantage of Nations* (New York, NY: The Free Press, 1990).

[50] Romer, Paul, "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94:5 (1986), 1002-1037.

[51] Romer, Paul, "Endogenous Technological Change", *Journal of Political Economy* 98:5 (1990), S71-S102.

[52] Rosenthal, Stuart, and William Strange, "Geography, Industrial Organization, and Agglomeration", *Review of Economics and Statistics* 85:2 (2003), 377-393.

[53] Rosenthal, Stuart and William Strange, "The Determinants of Agglomeration", *Journal of Urban Economics* 50 (2001), 191-229.

[54] Saxenian, AnnaLee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Cambridge, MA: Harvard University Press, 1994).

[55] Saxenian, AnnaLee, with Yasuyuki Motoyama and Xiaohong Quan, *Local and Global Networks of Immigrant Professionals in Silicon Valley* (San Francisco, CA: Public Policy Institute of California, 2002a).

[56] Saxenian, AnnaLee, "Silicon Valley's New Immigrant High-Growth Entrepreneurs", *Economic Development Quarterly* 16:1 (2002b), 20-31.

[57] Scherer, Frederic, "Using Linked Patent Data and R&D Data to Measure Technology Flows", in Griliches, Zvi (ed.) *R & D, Patents and Productivity* (Chicago, IL: University of Chicago Press, 1984).

[58] Silverman, Brian, "Technological Resources and the Direction of Corporate Diversification: Toward an Integration of the Resource-Based View and Transaction Cost Economics", *Management Science* 45:8 (1999), 1109-1124.

[59] Stephan, Paula, and Sharon Levin, "Exceptional Contributions to US Science by the Foreign-Born and Foreign-Educated", *Population Research and Policy Review* 20:1 (2001), 59-79.

[60] Streeter, Joanne, "Major Declines in Admissions of Immigrant Scientists and Engineers in Fiscal Year 1994", NSF SRS Issue Brief (1997).

[61] Thompson, Peter, and Melanie Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment", *American Economic Review* 95:1 (2005), 450-460.

[62] Wadhwa, Vivek, AnnaLee Saxenian, Ben Rissing, and Gary Gereffi, "America's New Immigrant Entrepreneurs I", Working Paper (2007).

Fig. 1: Ethnic Share of US Domestic Patents

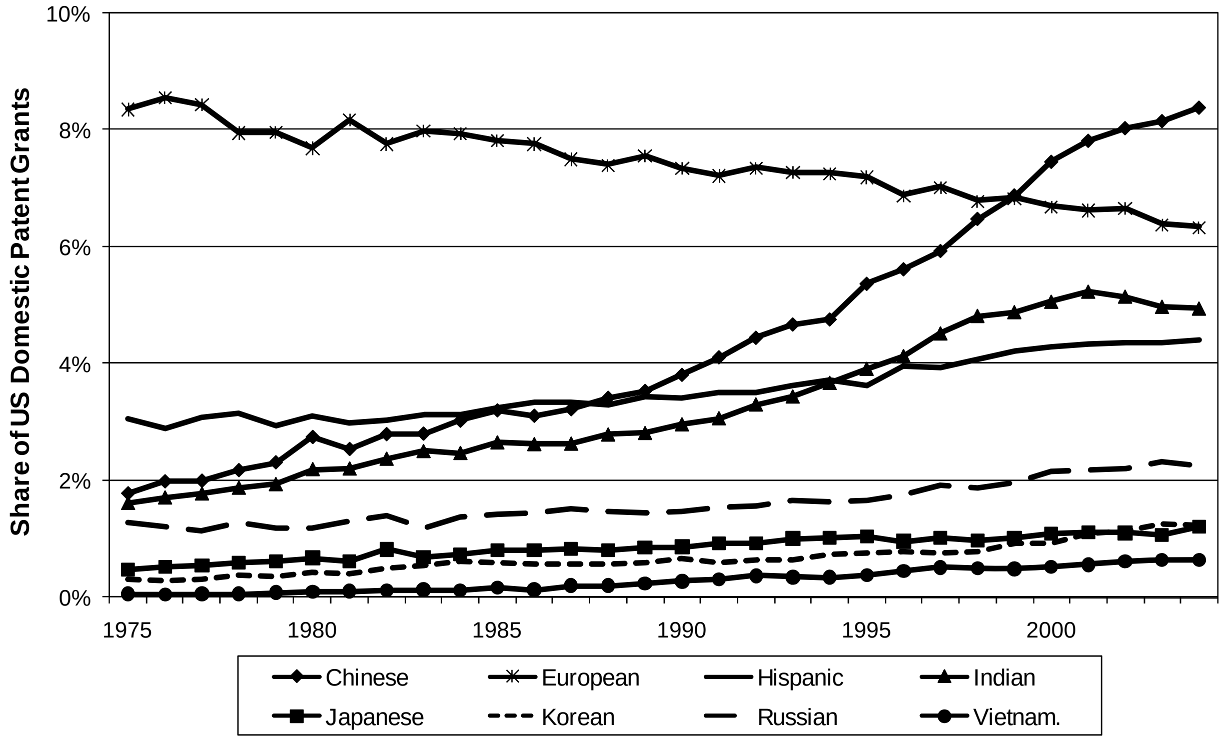

Fig. 2: HHI Concentration of US Patents

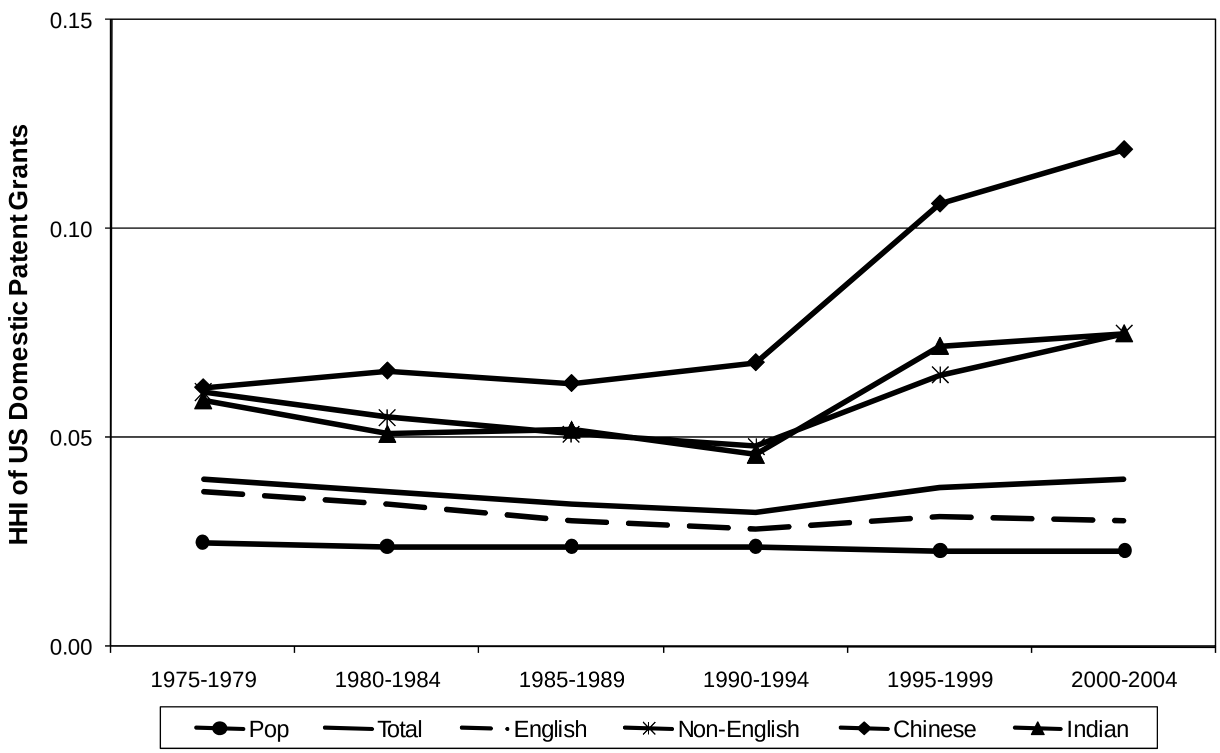

Fig. 3: EG Concentration of US Patents

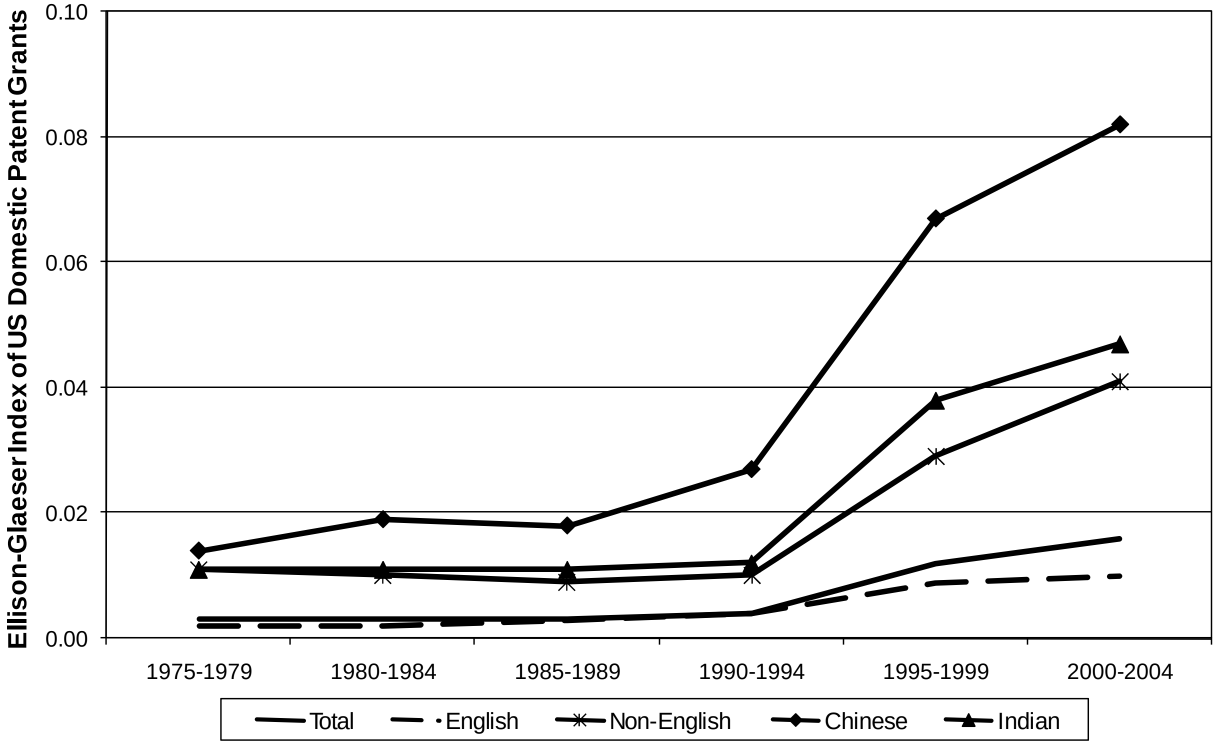

Fig. 4: Total US Ethnic Share by Technology

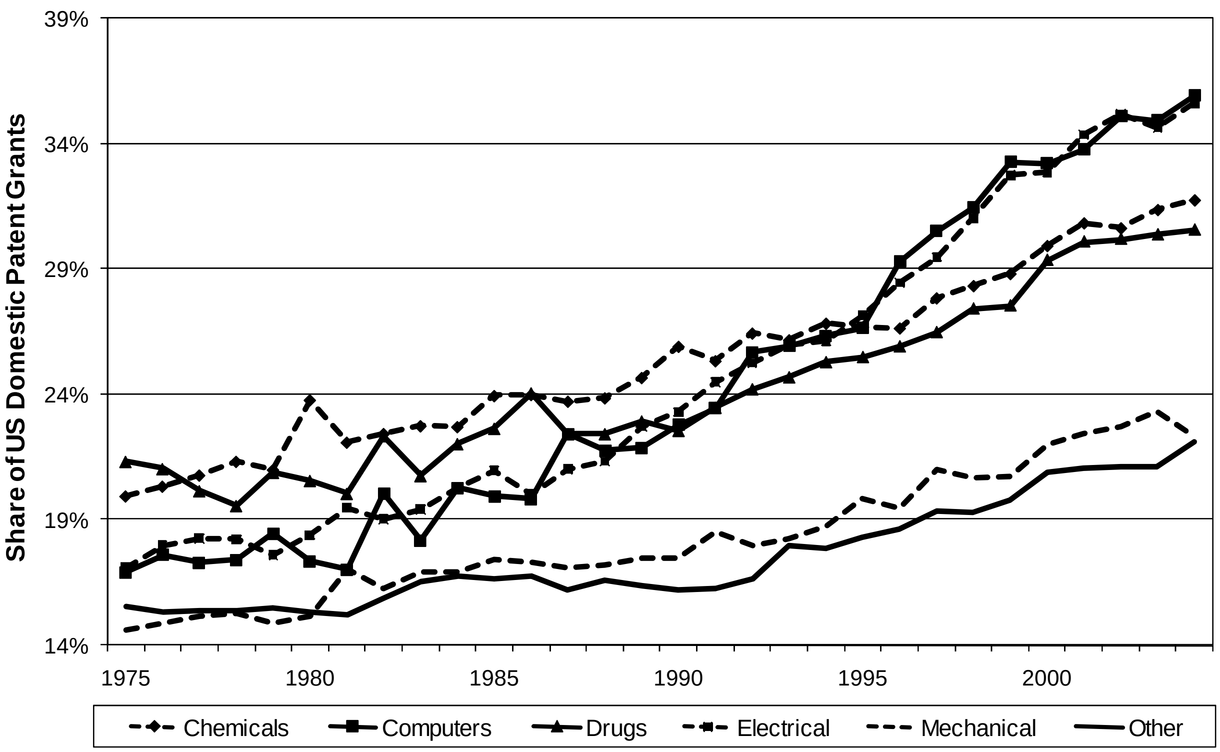

Fig. 5: Chinese Contribution by Technology

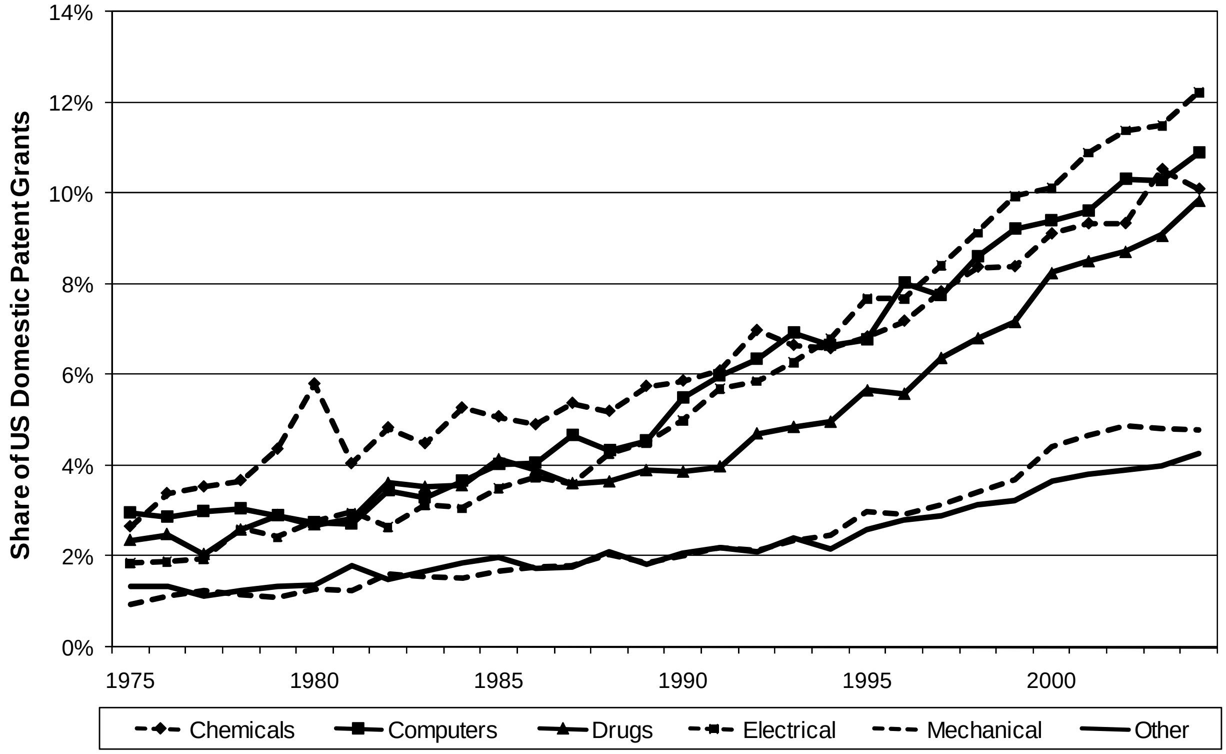

Fig. 6: Indian Contribution by Technology

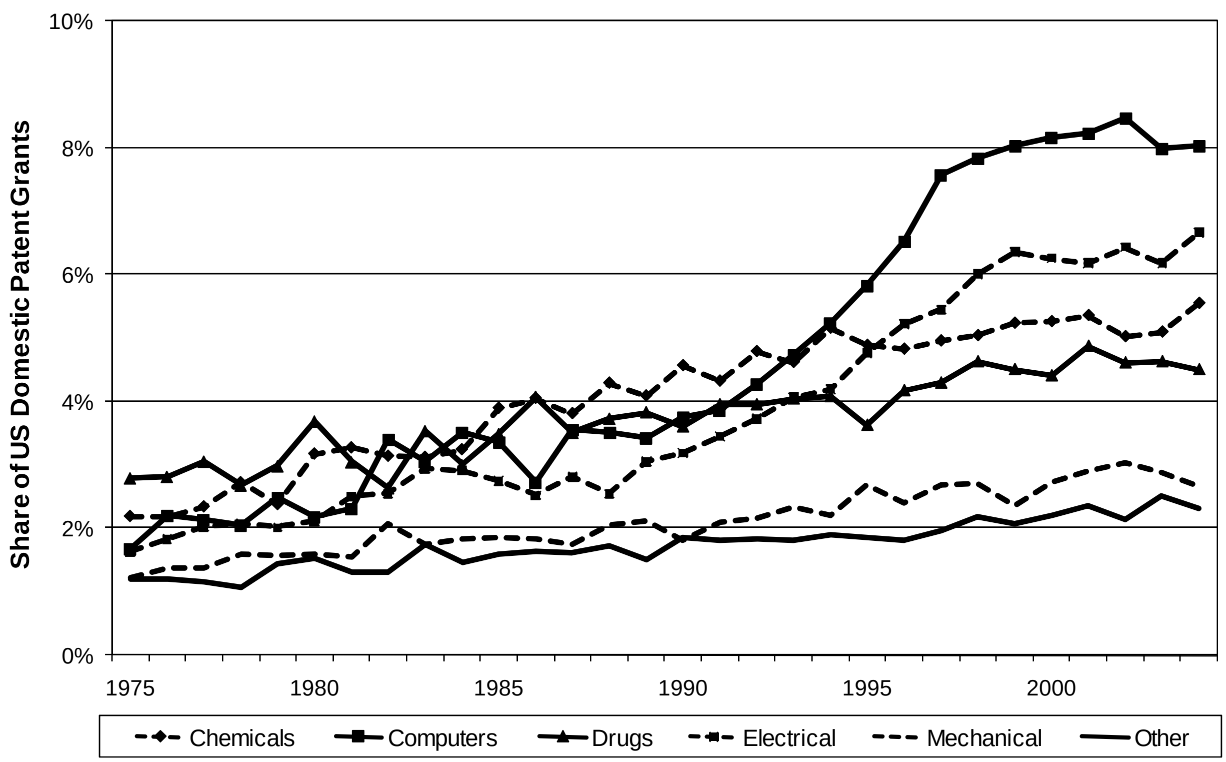

Fig. 7: Ethnic Concentration in Computers

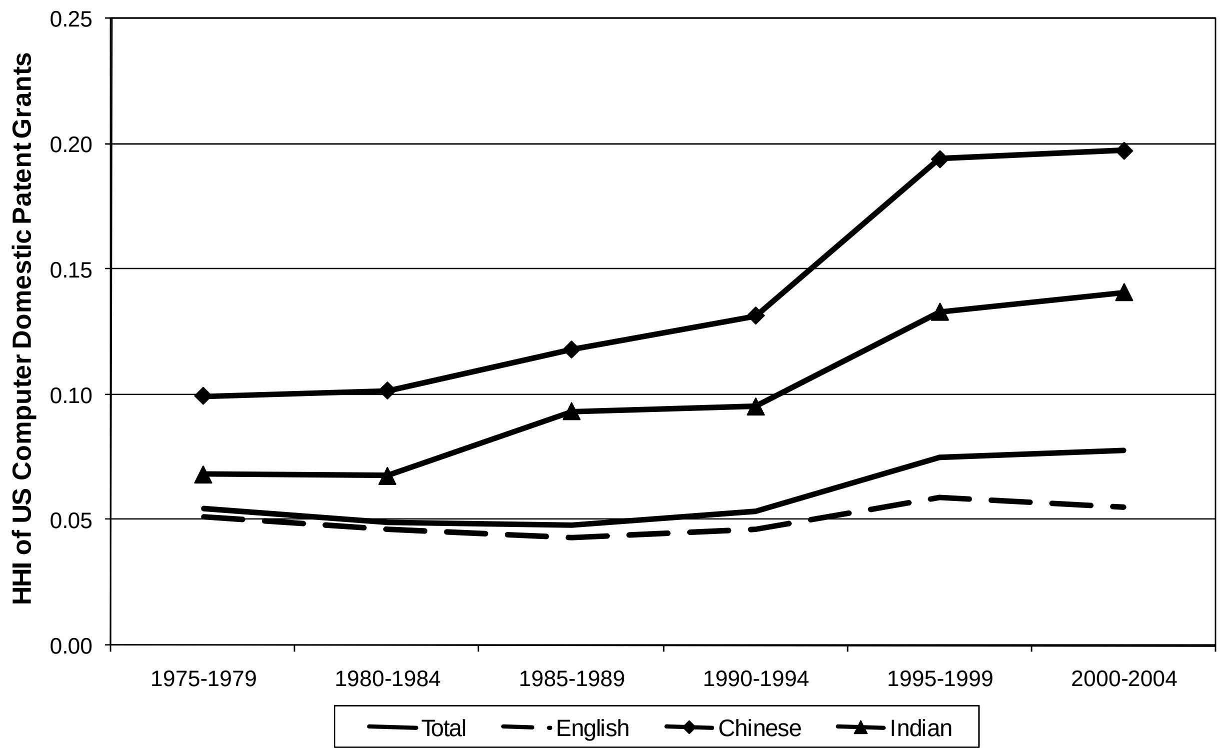

Fig. 8: Total US Ethnic Share by Institution

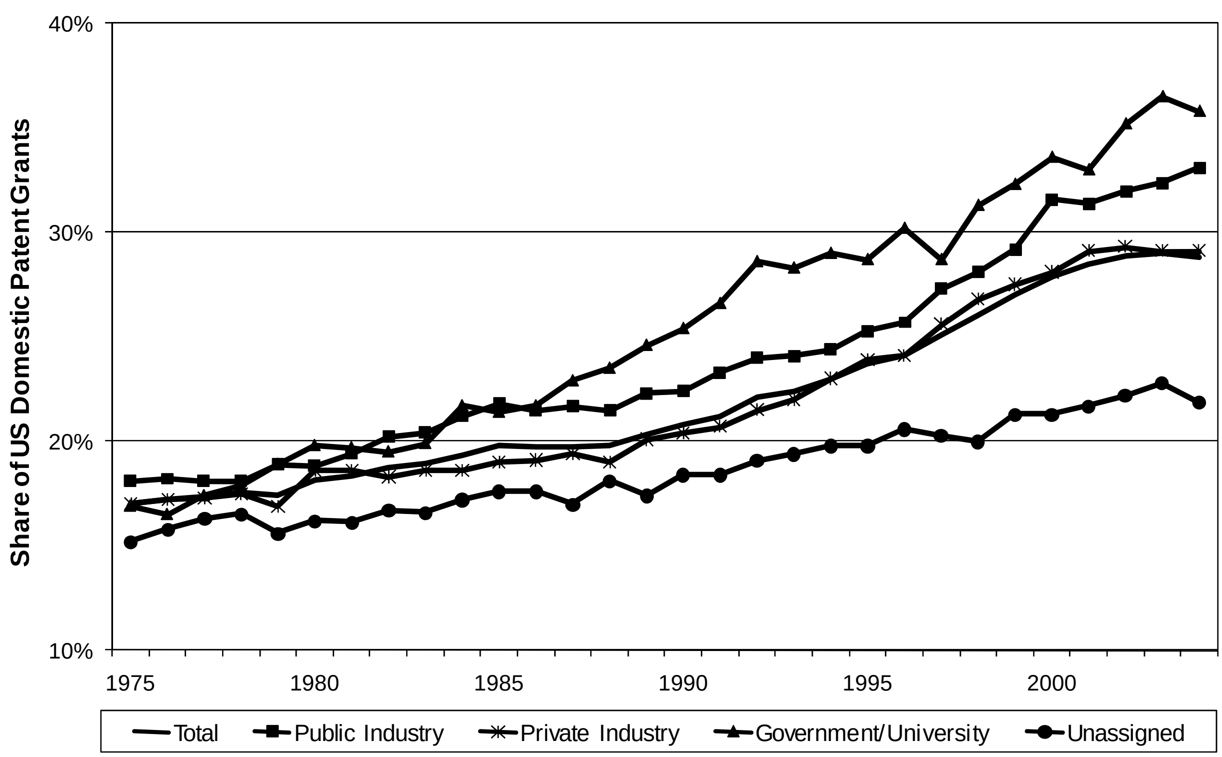

Fig. 9: Ethnic HHI, All Inventors

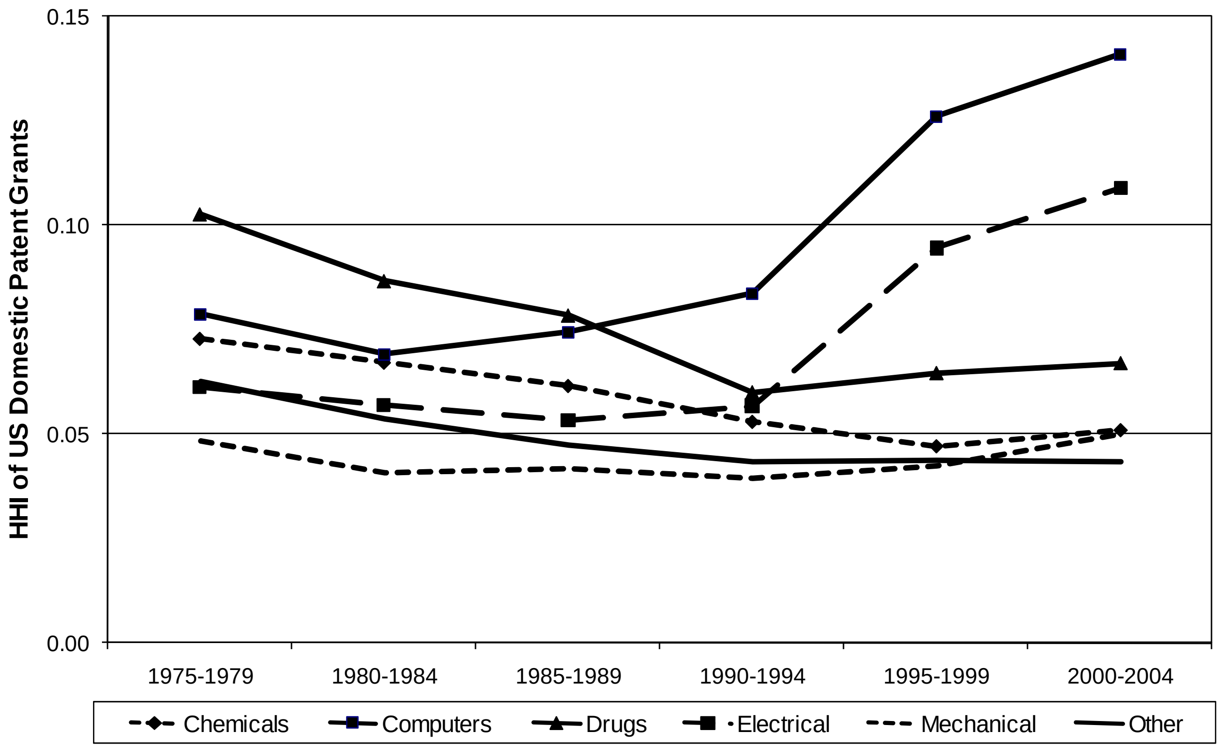

Fig. 10: Ethnic HHI, University & Government

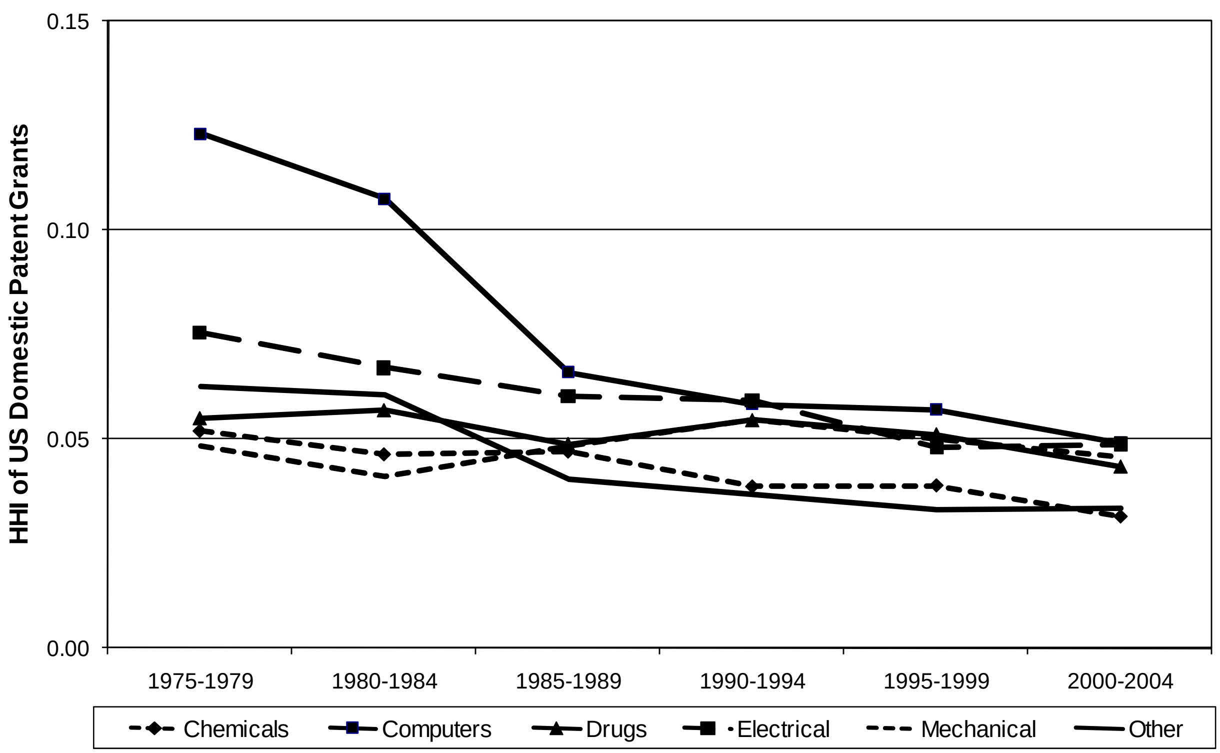

Table 1: Descriptive Statistics for Inventors Residing in US

	Ethnicity of Inventor								
	English	Chinese	European	Hispanic	Indian	Japanese	Korean	Russian	Vietnam.
	A. Ethnic Inventor Shares Estimated from US Inventor Records, 1975-2004								
1975-1979	82.5%	2.2%	8.3%	2.9%	1.9%	0.6%	0.3%	1.2%	0.1%
1980-1984	81.1%	2.9%	7.9%	3.0%	2.4%	0.7%	0.5%	1.3%	0.1%
1985-1989	79.8%	3.6%	7.5%	3.2%	2.9%	0.8%	0.6%	1.4%	0.2%
1990-1994	77.6%	4.6%	7.2%	3.5%	3.6%	0.9%	0.7%	1.5%	0.4%
1995-1999	73.9%	6.5%	6.8%	3.9%	4.8%	0.9%	0.8%	1.8%	0.5%
2000-2004	70.4%	8.5%	6.4%	4.2%	5.4%	1.0%	1.1%	2.2%	0.6%
Chemicals	73.4%	7.2%	7.5%	3.6%	4.5%	1.0%	0.8%	1.7%	0.3%
Computers	70.1%	8.2%	6.3%	3.8%	6.9%	1.1%	0.9%	2.1%	0.7%
Pharmaceuticals	72.9%	7.1%	7.4%	4.3%	4.2%	1.1%	0.9%	1.8%	0.4%
Electrical	71.6%	8.0%	6.8%	3.7%	4.9%	1.1%	1.1%	2.1%	0.7%
Mechanical	80.4%	3.2%	7.1%	3.5%	2.6%	0.7%	0.6%	1.6%	0.2%
Miscellaneous	81.3%	2.9%	7.0%	3.8%	2.1%	0.6%	0.6%	1.4%	0.3%
Top MSAs as a	KC (89)	SF (13)	NOR (12)	MIA (16)	SF (7)	SD (2)	BAL (2)	BOS (3)	AUS (2)
Percentage of	WS (88)	LA (8)	STL (11)	SA (9)	AUS (7)	SF (2)	LA (2)	NYC (3)	SF (1)
MSA's Patents	NAS (88)	AUS (6)	NYC (11)	WPB (7)	PRT (6)	LA (2)	SF (1)	SF (3)	LA (1)
	B. Ethnic Scientist and Engineer Shares Estimated from 1990 US Census Records								
Bachelors Share	87.6%	2.7%	2.3%	2.4%	2.3%	0.6%	0.5%	0.4%	1.2%
Masters Share	78.9%	6.7%	3.4%	2.2%	5.4%	0.9%	0.7%	0.8%	1.0%
Doctorate Share	71.2%	13.2%	4.0%	1.7%	6.5%	0.9%	1.5%	0.5%	0.4%

Notes: Panel A presents descriptive statistics for inventors residing in the US at the time of patent application. Inventor ethnicities are estimated through inventors' names using techniques described in the text. Patents are grouped by application years and major technology fields. Metropolitan Statistical Areas include AUS (Austin), BAL (Baltimore), BOS (Boston), KC (Kansas City), LA (Los Angeles), MIA (Miami), NAS (Nashville), NOR (New Orleans), NYC (New York City), PRT (Portland), SA (San Antonio), SD (San Diego), SF (San Francisco), STL (St. Louis), WPB (West Palm Beach), and WS (Winston-Salem). MSAs are identified from inventors' city names using city lists collected from the Office of Social and Economic Data Analysis at the University of Missouri, with a matching rate of 99%. Manual recoding further ensures all patents with more than 100 citations and all city names with more than 100 patents are identified. Panel B presents comparable statistics calculated from the 1990 Census using country of birth for scientists and engineers. Country groupings follow Table A3; English provides a residual in the Census statistics.

Table 2: Ethnic Inventor Contributions by MSA

	Total Patenting Share				non-English Ethnic Patenting Share				Chinese and Indian Patenting Share			
	1975-1984	1985-1994	1995-2004	2001-2006 (A)	1975-1984	1985-1994	1995-2004	2001-2006 (A)	1975-1984	1985-1994	1995-2004	2001-2006 (A)
Atlanta, GA	0.6%	1.0%	1.3%	1.5%	0.3%	0.7%	1.0%	1.1%	0.3%	0.7%	1.0%	1.2%
Austin, TX	0.4%	0.9%	1.8%	2.0%	0.5%	1.2%	1.9%	2.0%	0.4%	1.6%	2.3%	2.3%
Baltimore, MD	0.8%	0.8%	0.7%	0.7%	0.7%	0.7%	0.6%	0.5%	0.4%	0.5%	0.6%	0.5%
Boston, MA	3.6%	3.8%	3.9%	4.6%	3.9%	4.2%	4.1%	4.8%	4.0%	4.0%	3.6%	4.3%
Buffalo, NY	0.6%	0.5%	0.4%	0.3%	0.8%	0.6%	0.4%	0.3%	1.1%	0.7%	0.4%	0.3%
Charlotte, NC	0.3%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%	0.1%	0.2%	0.1%	0.2%
Chicago, IL	6.0%	4.6%	3.5%	3.2%	6.9%	5.0%	3.5%	3.0%	5.6%	3.9%	2.9%	2.8%
Cincinnati, OH	1.0%	1.1%	1.0%	1.0%	0.9%	0.9%	0.7%	0.7%	0.7%	1.0%	0.6%	0.6%
Cleveland, OH	2.3%	1.7%	1.3%	1.1%	2.5%	1.5%	1.0%	0.8%	2.5%	1.4%	0.9%	0.6%
Columbus, OH	0.7%	0.5%	0.5%	0.4%	0.6%	0.6%	0.4%	0.3%	0.8%	0.7%	0.3%	0.3%
Dallas-Fort Worth, TX	1.6%	2.0%	2.3%	2.1%	1.1%	1.9%	2.3%	2.2%	1.5%	2.4%	2.9%	2.8%
Denver, CO	1.0%	1.2%	1.3%	1.3%	0.8%	1.0%	0.9%	0.8%	0.8%	1.0%	0.6%	0.5%
Detroit, MI	3.1%	3.3%	2.9%	2.8%	3.1%	3.1%	2.6%	2.6%	3.2%	2.8%	2.5%	2.5%
Greensboro-W.S., NC	0.2%	0.3%	0.3%	0.2%	0.1%	0.2%	0.2%	0.1%	0.2%	0.2%	0.1%	0.1%
Hartford, CT	0.9%	0.9%	0.6%	0.6%	1.0%	0.8%	0.5%	0.5%	0.8%	0.6%	0.3%	0.4%
Houston, TX	2.3%	2.5%	1.9%	2.0%	1.8%	2.3%	1.8%	1.9%	2.2%	2.8%	1.8%	1.9%
Indianapolis, IN	0.8%	0.7%	0.7%	0.5%	0.6%	0.4%	0.4%	0.3%	0.7%	0.5%	0.4%	0.3%
Jacksonville, NC	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
Kansas City, MO	0.4%	0.3%	0.4%	0.3%	0.2%	0.2%	0.2%	0.2%	0.2%	0.1%	0.2%	0.2%
Las Vegas, NV	0.1%	0.1%	0.2%	0.3%	0.1%	0.1%	0.2%	0.2%	0.0%	0.1%	0.1%	0.1%
Los Angeles, CA	6.6%	6.1%	6.0%	5.7%	7.2%	7.2%	7.9%	7.3%	6.7%	6.9%	7.5%	7.0%
Memphis, TN	0.1%	0.2%	0.2%	0.3%	0.1%	0.1%	0.1%	0.2%	0.1%	0.1%	0.1%	0.1%
Miami, FL	0.8%	0.9%	0.7%	0.7%	1.0%	1.3%	1.0%	0.9%	0.5%	0.6%	0.5%	0.4%
Milwaukee, WI	1.0%	0.9%	0.8%	0.7%	0.8%	0.8%	0.6%	0.5%	0.5%	0.4%	0.5%	0.4%
Minneap.-St. Paul, MN	1.9%	2.4%	2.7%	2.8%	1.6%	2.0%	2.0%	2.0%	1.5%	1.7%	1.7%	1.8%

Table 2: Ethnic Inventor Contributions by MSA, continued

	Total Patenting Share				non-English Ethnic Patenting Share				Chinese and Indian Patenting Share			
	1975-1984	1985-1994	1995-2004	2001-2006 (A)	1975-1984	1985-1994	1995-2004	2001-2006 (A)	1975-1984	1985-1994	1995-2004	2001-2006 (A)
Nashville, TN	0.1%	0.2%	0.2%	0.2%	0.0%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
New Orleans, LA	0.3%	0.2%	0.2%	0.1%	0.3%	0.3%	0.1%	0.1%	0.2%	0.2%	0.0%	0.0%
New York, NY	11.5%	8.9%	7.3%	6.9%	16.6%	13.1%	10.1%	8.9%	16.6%	13.3%	9.7%	9.0%
Norfolk-VA Beach, VA	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
Orlando, FL	0.2%	0.3%	0.3%	0.3%	0.1%	0.2%	0.3%	0.3%	0.1%	0.2%	0.3%	0.3%
Philadelphia, PA	4.6%	4.0%	2.7%	2.8%	5.6%	4.9%	2.8%	2.9%	6.2%	5.8%	2.8%	3.0%
Phoenix, AZ	1.0%	1.2%	1.4%	1.3%	0.6%	1.1%	1.3%	1.2%	0.4%	1.0%	1.4%	1.3%
Pittsburgh, PA	2.0%	1.3%	0.8%	0.7%	2.2%	1.4%	0.6%	0.5%	2.2%	1.3%	0.5%	0.5%
Portland, OR	0.5%	0.8%	1.4%	1.6%	0.3%	0.6%	1.4%	1.6%	0.2%	0.6%	1.7%	2.0%
Providence, RI	0.3%	0.3%	0.3%	0.2%	0.3%	0.4%	0.3%	0.2%	0.2%	0.3%	0.2%	0.2%
Raleigh-Durham, NC	0.3%	0.6%	1.1%	1.5%	0.3%	0.6%	1.0%	1.3%	0.3%	0.8%	1.0%	1.2%
Richmond, VA	0.3%	0.3%	0.2%	0.2%	0.3%	0.3%	0.2%	0.2%	0.3%	0.4%	0.2%	0.2%
Sacramento, CA	0.2%	0.4%	0.5%	0.5%	0.2%	0.4%	0.5%	0.5%	0.2%	0.3%	0.5%	0.5%
Salt Lake City, UT	0.4%	0.5%	0.6%	0.6%	0.2%	0.4%	0.3%	0.3%	0.2%	0.3%	0.3%	0.3%
San Antonio, TX	0.1%	0.2%	0.2%	0.2%	0.1%	0.2%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%
San Diego, CA	1.1%	1.6%	2.2%	2.8%	1.1%	1.6%	2.6%	3.6%	0.8%	1.4%	2.4%	3.9%
San Francisco, CA	4.8%	6.6%	12.1%	13.2%	6.2%	9.3%	19.3%	19.9%	8.4%	13.0%	25.4%	24.0%
Seattle, WA	0.9%	1.3%	1.9%	3.4%	0.8%	1.1%	1.8%	3.5%	0.6%	1.0%	1.8%	3.7%
St. Louis, MO	1.0%	0.9%	0.8%	0.8%	0.9%	0.8%	0.8%	0.7%	1.0%	0.8%	0.4%	0.4%
Tallahassee, FL	0.4%	0.5%	0.4%	0.4%	0.3%	0.4%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%
Washington, DC	1.5%	1.5%	1.4%	1.6%	1.6%	1.6%	1.5%	1.7%	1.6%	1.7%	1.5%	1.7%
West Palm Beach, FL	0.3%	0.5%	0.4%	0.4%	0.3%	0.5%	0.4%	0.4%	0.3%	0.3%	0.2%	0.2%
Other 234 MSAs	21.8%	22.3%	20.7%	18.4%	18.1%	18.1%	15.6%	13.6%	19.7%	18.2%	14.6%	12.7%
Not in an MSA	9.0%	8.2%	6.6%	6.2%	6.3%	5.4%	3.7%	4.1%	5.2%	3.8%	2.5%	2.7%

Notes: See Table 1. The first three columns of each grouping are for granted patents. The fourth column, marked with (A), is for published patent applications.

Table 3: Ethnic Inventors and MSA Characteristics, Weighted Estimations

	English		Chinese		Indian		European		Hispanic	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent Variable is Share of 1985-2004 Ethnic Patenting in the MSA									
1975-1984 Share of Ethnic Patents in MSA		0.842 (0.284)		0.865 (0.501)		0.796 (0.186)		0.646 (0.053)		0.526 (0.185)
Log Population of MSA	0.573 (0.076)	-0.132 (0.260)	0.475 (0.099)	-0.273 (0.495)	0.457 (0.199)	-0.176 (0.186)	0.650 (0.191)	0.117 (0.066)	0.812 (0.071)	0.268 (0.200)
Log Population Density of MSA	0.251 (0.105)	-0.063 (0.134)	-0.140 (0.129)	-0.253 (0.166)	0.143 (0.238)	-0.223 (0.146)	0.329 (0.211)	-0.004 (0.084)	-0.080 (0.106)	-0.100 (0.078)
Coastal Access of MSA	0.029 (0.137)	0.177 (0.161)	0.378 (0.266)	0.294 (0.160)	0.240 (0.237)	0.327 (0.221)	0.063 (0.146)	0.190 (0.132)	0.331 (0.135)	0.269 (0.106)
Share of Population with Bachelors Ed.	0.429 (0.257)	0.268 (0.163)	0.505 (0.399)	0.184 (0.163)	0.602 (0.378)	0.353 (0.253)	0.498 (0.270)	0.301 (0.201)	0.303 (0.216)	0.220 (0.174)
Share of Population under 30 in Age	-0.779 (0.566)	-0.711 (0.456)	-1.320 (1.150)	-1.031 (0.684)	-1.291 (0.980)	-1.161 (0.824)	-0.641 (0.569)	-0.667 (0.519)	-0.558 (0.535)	-0.581 (0.493)
Share of Population over 60 in Age	-0.452 (0.347)	-0.567 (0.325)	-0.757 (0.704)	-0.804 (0.535)	-0.703 (0.598)	-0.844 (0.549)	-0.175 (0.362)	-0.432 (0.326)	-0.275 (0.334)	-0.400 (0.327)
Share of Population Female	-0.313 (0.256)	-0.451 (0.268)	-0.576 (0.516)	-0.968 (0.592)	-0.090 (0.485)	-0.632 (0.489)	0.155 (0.340)	-0.295 (0.251)	-0.128 (0.247)	-0.375 (0.285)
R-Squared	0.84	0.88	0.54	0.69	0.61	0.74	0.82	0.91	0.90	0.92

Notes: Estimations provide partial correlations for ethnic patenting undertaken in 244 MSAs over the 1985-2004 period. The dependent variable is the MSA's share of indicated ethnic invention relative to the MSA sample. Explanatory regressors are from the 1990 Census of Populations, excepting coastal access and the lagged ethnic patenting share. The latter is ethnic specific and is calculated for the 1975-1984 pre-period from the ethnic patenting database. Estimations are weighted by MSA populations. Variables are transformed to unit standard deviation for interpretation. Robust standard errors are reported in parenthesis.

Table 4: Ethnic Inventors and MSA Characteristics, Unweighted Estimations

	English		Chinese		Indian		European		Hispanic	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent Variable is Share of 1985-2004 Ethnic Patenting in the MSA									
1975-1984 Share of		0.884		0.968		0.726		0.643		0.655
Ethnic Patents in MSA		(0.255)		(0.586)		(0.262)		(0.107)		(0.271)
Log Population	0.810	-0.029	0.647	-0.230	0.684	0.037	0.845	0.261	0.901	0.250
of MSA	(0.106)	(0.171)	(0.145)	(0.431)	(0.185)	(0.134)	(0.166)	(0.102)	(0.075)	(0.189)
Log Population	0.053	0.026	-0.047	-0.019	-0.002	-0.018	0.020	0.016	-0.043	-0.003
Density of MSA	(0.034)	(0.026)	(0.029)	(0.039)	(0.051)	(0.030)	(0.050)	(0.020)	(0.023)	(0.015)
Coastal Access	-0.027	0.022	0.052	0.067	0.012	0.046	-0.009	0.020	0.054	0.043
of MSA	(0.035)	(0.039)	(0.057)	(0.047)	(0.050)	(0.055)	(0.033)	(0.030)	(0.033)	(0.026)
Share of Population	0.123	0.091	0.084	0.041	0.113	0.087	0.094	0.080	0.070	0.067
with Bachelors Ed.	(0.034)	(0.023)	(0.050)	(0.025)	(0.048)	(0.035)	(0.034)	(0.026)	(0.029)	(0.025)
Share of Population	-0.151	-0.145	-0.115	-0.152	-0.139	-0.150	-0.078	-0.110	-0.045	-0.090
under 30 in Age	(0.064)	(0.056)	(0.111)	(0.104)	(0.100)	(0.091)	(0.065)	(0.055)	(0.056)	(0.061)
Share of Population	-0.102	-0.135	-0.078	-0.151	-0.086	-0.140	-0.015	-0.081	-0.012	-0.076
over 60 in Age	(0.051)	(0.053)	(0.086)	(0.103)	(0.078)	(0.084)	(0.053)	(0.045)	(0.047)	(0.056)
Share of Population	-0.056	-0.050	-0.055	-0.058	-0.055	-0.057	-0.032	-0.039	-0.033	-0.042
Female	(0.023)	(0.021)	(0.037)	(0.033)	(0.033)	(0.033)	(0.021)	(0.019)	(0.021)	(0.021)
R-Squared	0.79	0.85	0.45	0.65	0.54	0.64	0.78	0.86	0.83	0.87

Notes: See Table 3. Estimations are unweighted.

Table 5: Concentration Ratios of Invention

	Total Population	Total Invention	English Invention	Non-Eng. Invention	Chinese Invention	Indian Invention
			A. Herfindahl-Hirschman Index			
1975-1979	0.025	0.040	0.037	0.061	0.062	0.059
1980-1984	0.024	0.037	0.034	0.055	0.066	0.051
1985-1989	0.024	0.034	0.030	0.051	0.063	0.052
1990-1994	0.024	0.032	0.028	0.048	0.068	0.046
1995-1999	0.023	0.038	0.031	0.065	0.106	0.072
2000-2004	0.023	0.040	0.030	0.075	0.119	0.075
Mean	0.024	0.037	0.032	0.059	0.081	0.059
			B. Share in Top 5 MSAs from 1975-1984			
1975-1979	28.2%	37.8%	35.9%	46.7%	48.0%	43.4%
1980-1984	27.5%	35.7%	33.8%	44.0%	49.5%	40.1%
1985-1989	27.4%	33.7%	31.4%	43.0%	49.2%	41.2%
1990-1994	27.1%	32.2%	29.6%	41.2%	48.6%	38.5%
1995-1999	26.5%	33.7%	29.8%	44.6%	53.3%	43.3%
2000-2004	26.5%	33.1%	28.0%	45.1%	53.8%	41.6%
Mean	27.2%	34.4%	31.4%	44.1%	50.4%	41.4%
			C. Ellison-Glaeser Index Relative to MSA Populations			
1975-1979	n.a.	0.003	0.002	0.011	0.014	0.011
1980-1984		0.003	0.002	0.010	0.019	0.011
1985-1989		0.003	0.003	0.009	0.018	0.011
1990-1994		0.004	0.004	0.010	0.027	0.012
1995-1999		0.012	0.009	0.029	0.067	0.038
2000-2004		0.016	0.010	0.041	0.082	0.047
Mean		0.007	0.005	0.018	0.038	0.022

Notes: Metrics consider agglomeration of US domestic invention across 283 MSAs, with invention in rural areas excluded. Top 5 MSAs are kept constant from 1975-1984 rankings: New York City, Los Angeles, Chicago, Philadelphia, and San Francisco. Ellison and Glaeser metrics consider agglomeration of invention relative to MSA populations. These latter metrics abstract from plant Herfindahl corrections. General population counts from 1995-1999 are used for 2000-2004.

Table 6: Concentration Ratios of Invention by Technology Group

	Chemicals	Computers & Comm.	Drugs & Medical	Electrical & Electronic	Mechanical	Miscellaneous
	A. Herfindahl-Hirschman Index for All Patents Within Technology Group					
1975-1979	0.053	0.055	0.070	0.043	0.032	0.039
1980-1984	0.048	0.050	0.061	0.039	0.030	0.035
1985-1989	0.043	0.048	0.055	0.036	0.029	0.031
1990-1994	0.038	0.054	0.047	0.037	0.028	0.028
1995-1999	0.033	0.075	0.050	0.052	0.029	0.027
2000-2004	0.034	0.078	0.053	0.059	0.032	0.026
Mean	0.041	0.060	0.056	0.044	0.030	0.031
	B. HHI for English Patents Within Technology Group					
1975-1979	0.049	0.051	0.063	0.040	0.030	0.036
1980-1984	0.043	0.046	0.056	0.035	0.028	0.032
1985-1989	0.038	0.043	0.050	0.033	0.027	0.028
1990-1994	0.033	0.046	0.044	0.032	0.026	0.025
1995-1999	0.029	0.059	0.046	0.038	0.026	0.023
2000-2004	0.028	0.055	0.048	0.040	0.028	0.022
Mean	0.037	0.050	0.051	0.036	0.028	0.028
	C. HHI for non-English Patents Within Technology Group					
1975-1979	0.073	0.079	0.103	0.061	0.048	0.062
1980-1984	0.067	0.069	0.087	0.057	0.041	0.053
1985-1989	0.062	0.074	0.078	0.053	0.042	0.047
1990-1994	0.053	0.084	0.060	0.057	0.039	0.043
1995-1999	0.047	0.126	0.065	0.095	0.042	0.044
2000-2004	0.051	0.141	0.067	0.109	0.050	0.043
Mean	0.059	0.095	0.077	0.072	0.044	0.049

Notes: See Table 5. Patents are grouped into the major technology categories given in the column headers.

Table 7: Concentration Ratios of Invention by Institution

	Chemicals	Computers & Comm.	Drugs & Medical	Electrical & Electronic	Mechanical	Miscellaneous
			A. Herfindahl-Hirschman Index for All Industry Patents			
1975-1979	0.058	0.056	0.086	0.044	0.033	0.040
1980-1984	0.053	0.050	0.076	0.040	0.031	0.037
1985-1989	0.047	0.050	0.064	0.036	0.030	0.030
1990-1994	0.042	0.056	0.054	0.038	0.031	0.027
1995-1999	0.035	0.080	0.058	0.055	0.031	0.025
2000-2004	0.037	0.082	0.061	0.064	0.037	0.025
Mean	0.045	0.062	0.066	0.046	0.032	0.031
			B. HHI for All University and Government Patents			
1975-1979	0.043	0.088	0.043	0.054	0.041	0.040
1980-1984	0.039	0.068	0.046	0.050	0.039	0.040
1985-1989	0.036	0.059	0.044	0.046	0.041	0.029
1990-1994	0.033	0.049	0.047	0.052	0.040	0.031
1995-1999	0.035	0.048	0.041	0.045	0.040	0.027
2000-2004	0.033	0.044	0.038	0.042	0.039	0.029
Mean	0.036	0.059	0.043	0.048	0.040	0.033
			C. HHI for non-English Industry Patents			
1975-1979	0.078	0.079	0.118	0.061	0.046	0.061
1980-1984	0.072	0.068	0.110	0.057	0.042	0.052
1985-1989	0.067	0.078	0.091	0.053	0.042	0.045
1990-1994	0.058	0.089	0.071	0.060	0.041	0.038
1995-1999	0.050	0.133	0.076	0.103	0.044	0.038
2000-2004	0.056	0.148	0.077	0.118	0.055	0.038
Mean	0.064	0.099	0.091	0.075	0.045	0.045
			D. HHI for non-English University and Government Patents			
1975-1979	0.052	0.123	0.055	0.075	0.048	0.063
1980-1984	0.046	0.108	0.057	0.067	0.041	0.060
1985-1989	0.047	0.066	0.049	0.060	0.048	0.040
1990-1994	0.039	0.058	0.055	0.059	0.055	0.037
1995-1999	0.039	0.057	0.051	0.048	0.050	0.033
2000-2004	0.031	0.049	0.043	0.049	0.046	0.034
Mean	0.042	0.077	0.052	0.060	0.048	0.044

Notes: See Table 5. Patents are grouped into the major technology categories given in the column headers.

Table A1: Coagglomeration of US Ethnic Invention

	Chinese	English	European	Hispanic	Indian	Japanese	Korean	Russian	Vietnam.
					A. 1975-1979 Coagglomeration of Ethnic Invention				
Chinese	0.014								
English	0.004	0.002							
European	0.011	0.004	0.014						
Hispanic	0.010	0.003	0.009	0.011					
Indian	0.011	0.004	0.012	0.009	0.011				
Japanese	0.010	0.005	0.005	0.011	0.005	0.034			
Korean	0.009	0.004	0.009	0.008	0.008	0.012	0.012		
Russian	0.011	0.005	0.012	0.011	0.011	0.012	0.010	0.015	
Vietnam.	0.011	0.004	0.009	0.012	0.008	0.020	0.010	0.013	0.024
					B. 2000-2004 Coagglomeration of Ethnic Invention				
Chinese	0.082								
English	0.024	0.010							
European	0.033	0.011	0.016						
Hispanic	0.034	0.010	0.014	0.016					
Indian	0.059	0.019	0.025	0.025	0.047				
Japanese	0.082	0.024	0.032	0.034	0.058	0.084			
Korean	0.075	0.020	0.030	0.031	0.053	0.075	0.071		
Russian	0.051	0.015	0.022	0.022	0.037	0.051	0.048	0.034	
Vietnam.	0.086	0.026	0.033	0.035	0.062	0.087	0.078	0.051	0.097

Notes: Metrics consider coagglomeration of ethnic invention relative to MSA populations.

Table A2: Concentration Ratios at Sub-Category Levels

	Chemicals	Computers & Comm.	Drugs & Medical	Electrical & Electronic	Mechanical	Miscellaneous
	A. Herfindahl-Hirschman Index for All Patents Within Technology Group					
1975-1979	0.053	0.055	0.070	0.043	0.032	0.039
1980-1984	0.048	0.050	0.061	0.039	0.030	0.035
1985-1989	0.043	0.048	0.055	0.036	0.029	0.031
1990-1994	0.038	0.054	0.047	0.037	0.028	0.028
1995-1999	0.033	0.075	0.050	0.052	0.029	0.027
2000-2004	0.034	0.078	0.053	0.059	0.032	0.026
Mean	0.041	0.060	0.056	0.044	0.030	0.031
	B. Unweighted HHI Average Across Sub-Category Technology Groups					
1975-1979	0.057	0.059	0.072	0.051	0.044	0.052
1980-1984	0.053	0.059	0.069	0.048	0.040	0.050
1985-1989	0.050	0.064	0.063	0.046	0.042	0.042
1990-1994	0.041	0.073	0.054	0.046	0.049	0.040
1995-1999	0.039	0.095	0.057	0.057	0.048	0.041
2000-2004	0.040	0.102	0.062	0.060	0.049	0.051
Mean	0.047	0.075	0.063	0.051	0.045	0.046
	C. Weighted HHI Average Across Sub-Category Technology Groups					
1975-1979	0.060	0.059	0.083	0.047	0.038	0.047
1980-1984	0.053	0.055	0.071	0.044	0.035	0.044
1985-1989	0.047	0.055	0.066	0.043	0.036	0.038
1990-1994	0.041	0.062	0.054	0.045	0.040	0.035
1995-1999	0.037	0.085	0.058	0.064	0.041	0.035
2000-2004	0.038	0.088	0.062	0.072	0.047	0.042
Mean	0.046	0.068	0.066	0.052	0.040	0.040

Notes: See Table 6.

Table A3: Descriptive Statistics for Inventors Residing in Foreign Countries and Regions

Summary Statistics for Full and Restricted Matching Procedures

		Percentage of Region's Inventors Matched with Ethnic Database		Percentage of Region's Inventors Assigned Ethnicity of Their Region		Percentage of Region's Inventors Assigned Ethnicity of Region (Partial)	
	Obs.	Full	Restrict.	Full	Restrict.	Full	Restrict.
United Kingdom	187,266	99%	95%	85%	83%	92%	91%
China, Singapore	167,370	100%	98%	88%	89%	91%	91%
Western Europe	1,210,231	98%	79%	66%	46%	73%	58%
Hispanic Nations	27,298	99%	74%	74%	69%	93%	93%
India	13,582	93%	76%	88%	88%	90%	89%
Japan	1,822,253	100%	89%	100%	96%	100%	96%
South Korea	127,975	100%	100%	84%	83%	89%	88%
Russia	33,237	94%	78%	81%	84%	93%	94%
Vietnam	41	100%	98%	36%	43%	44%	43%

Complete Ethnic Composition of Region's Inventors (Full Matching)

	English	Chinese	European	Hispanic	Indian	Japanese	Korean	Russian	Vietnam.
United Kingdom	**85%**	2%	5%	3%	2%	0%	0%	2%	0%
China, Singapore	3%	**88%**	1%	1%	1%	1%	4%	1%	1%
Western Europe	21%	1%	**66%**	8%	1%	0%	0%	3%	0%
Hispanic Nations	11%	1%	10%	**74%**	0%	1%	0%	2%	0%
India	3%	1%	1%	5%	**88%**	0%	0%	2%	0%
Japan	0%	0%	0%	0%	0%	**100%**	0%	0%	0%
South Korea	2%	11%	0%	1%	0%	1%	**84%**	1%	0%
Russia	5%	1%	3%	9%	0%	0%	0%	**81%**	0%
Vietnam	17%	21%	12%	0%	0%	10%	2%	2%	**36%**

Notes: Matching is undertaken at inventor level using the Full and Restricted Matching procedures outlined in the text. The middle columns of the top panel summarize the share of each region's inventors assigned the ethnicity of that region; the complete composition for the Full Matching procedure is detailed in the bottom panel. The right-hand columns in the top panel document the percentage of the region's inventors assigned at least partially to their region's ethnicity.

Greater China includes Mainland China, Hong Kong, Macao, and Taiwan. Western Europe includes Austria, Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, Netherlands, Norway, Poland, Sweden, and Switzerland. Hispanic Nations includes Argentina, Belize, Brazil, Chile, Columbia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Philippines, Portugal, Spain, Uruguay, and Venezuela. Russia includes former Soviet Union countries.

Table A4: Most Common Ethnic Surnames for Inventors Residing in the US

Chinese		English		European		Hispanic / Filipino		Indian / Hindi	
CAI	585	ADAMS	4,490	ABEL	269	ACOSTA	171	ACHARYA	338
CAO	657	ALLEN	5,074	ALBRECHT	564	AGUILAR	138	AGARWAL	580
CHAN	3,096	ANDERSON	10,719	ANTOS	230	ALVAREZ	446	AGGARWAL	282
CHANG	3,842	BAILEY	2,431	AUERBACH	193	ANDREAS	128	AGRAWAL	797
CHAO	796	BAKER	4,671	BAER	422	AYER	166	AHMAD	355
CHAU	486	BELL	2,738	BAERLOCHER	252	AYRES	180	AHMED	652
CHEN	12,860	BENNETT	2,734	BAUER	1,470	BALES	240	AKRAM	640
CHENG	2,648	BROOKS	2,015	BECHTEL	179	BLANCO	141	ALI	559
CHEUNG	950	BROWN	11,662	BECK	1,712	BOLANOS	130	ARIMILLI	432
CHIANG	1,112	BURNS	2,098	BENDER	650	BOLES	118	ARORA	214
CHIEN	429	CAMPBELL	3,959	BERG	1,465	CABRAL	154	ASH	290
CHIN	423	CARLSON	2,745	BERGER	1,304	CABRERA	163	BALAKRISHNAN	228
CHIU	924	CARTER	2,658	BOEHM	256	CALDERON	124	BANERJEE	371
CHOU	1,144	CHANG	2,032	BOUTAGHOU	266	CASTANEDA	116	BASU	233
CHOW	1,139	CLARK	5,493	CARON	290	CASTILLO	124	BHAT	224
CHU	2,353	COHEN	2,626	CERAMI	172	CASTRO	119	BHATIA	411
DENG	439	COLE	2,143	CHANDRARATNA	229	CHAVEZ	194	BHATT	242
DING	589	COLLINS	2,992	CHEVALLIER	204	CONTRERAS	137	BHATTACHARYA	216
DONG	492	COOK	3,556	DIETRICH	312	CRUZ	319	BHATTACHARYYA	265
FAN	1,036	COOPER	3,045	DIETZ	496	CUEVAS	123	BOSE	238
FANG	846	COX	2,407	EBERHARDT	192	DAS	213	CHANDRA	221
FENG	658	DAVIS	8,848	EHRLICH	311	DELGADO	216	CHATTERJEE	647
FONG	727	EDWARDS	3,375	ERRICO	190	DIAS	174	DAOUD	305
FU	767	EVANS	4,082	FARKAS	169	DIAZ	584	DAS	522
FUNG	455	FISCHER	2,081	FERRARI	177	DOMINGUEZ	195	DATTA	424
GAO	785	FISHER	2,748	FISCHELL	280	DURAN	142	DE	234
GUO	921	FOSTER	2,616	FUCHS	394	ELIAS	230	DESAI	974
HAN	777	FOX	1,990	GAISER	193	ESTRADA	142	DIXIT	256
HE	1,159	GARDNER	2,412	GELARDI	176	FERNANDES	152	DUTTA	338
HO	1,282	GORDON	2,315	GRILLIOT	201	FERNANDEZ	546	GANDHI	228
HSIEH	980	GRAHAM	2,042	GUEGLER	179	FIGUEROA	146	GARG	345
HSU	3,034	GRAY	2,626	GUNTER	177	FLORES	191	GHOSH	661
HU	1,695	GREEN	3,540	GUNTHER	247	FREITAS	132	GOEL	279
HUANG	4,605	HALL	4,907	HAAS	843	GAGNON	265	GUPTA	1,935
HUI	451	HAMILTON	1,991	HAMPEL	187	GARCIA	1,310	HASSAN	217
HUNG	562	HANSON	2,148	HANSEN	2,947	GARZA	167	HUSSAIN	233
HWANG	800	HARRIS	4,793	HARTMAN	1,214	GOMES	199	HUSSAINI	299
JIANG	1,399	HAYES	2,031	HARTMANN	385	GOMEZ	413	ISLAM	266
KAO	714	HILL	3,590	HAUSE	266	GONSALVES	141	IYER	601
KUO	1,157	HOFFMAN	2,387	HECHT	245	GONZALES	281	JAIN	912
LAI	1,134	HOWARD	2,160	HEINZ	168	GONZALEZ	1,055	JOSHI	886
LAM	1,336	HUGHES	2,198	HORODYSKY	230	GUTIERREZ	601	KAMATH	219
LAU	1,320	JACKSON	3,980	HORVATH	387	GUZMAN	139	KAPOOR	222
LEE	4,006	JENSEN	2,361	IACOVELLI	287	HALASA	202	KHANNA	378
LEUNG	1,165	JOHNSON	17,960	JACOBS	1,962	HERNANDEZ	703	KRISHNAMURTHY	369
LEW	460	JONES	10,630	KARR	196	HERRERA	171	KRISHNAN	512
LI	6,863	KELLER	2,041	KASPER	227	HERRON	450	KULKARNI	299
LIANG	1,173	KELLY	2,775	KEMPF	228	HIDALGO	186	KUMAR	2,005
LIAO	553	KENNEDY	2,208	KNAPP	833	JIMENEZ	246	LAL	366
LIM	485	KING	4,686	KNIFTON	206	LEE	237	MALIK	532
LIN	5,770	KLEIN	2,347	KOENIG	521	LOPEZ	738	MATHUR	306
LING	521	LARSON	2,537	KRESGE	179	MACHADO	135	MEHROTRA	265

Table A4: Most Common US Ethnic Surnames (continued)

Chinese		English		European		Hispanic / Filipino		Indian / Hindi	
LIU	6,406	LEE	9,490	LANGE	757	MARIN	177	MEHTA	925
LO	1,053	LEWIS	4,732	LASKARIS	192	MARQUEZ	117	MENON	325
LU	2,289	LONG	2,392	LEMELSON	324	MARTIN	183	MISHRA	348
LUO	815	MARSHALL	2,088	LIOTTA	171	MARTINEZ	1,112	MISRA	282
MA	1,708	MARTIN	6,773	LORENZ	341	MATIS	249	MOOKHERJEE	272
MAO	545	MILLER	14,942	LUDWIG	500	MEDINA	192	MUKHERJEE	327
NG	1,132	MITCHELL	3,075	LUTZ	679	MENARD	149	MURTHY	236
ONG	473	MOORE	6,459	MAIER	492	MENDOZA	173	NAGARAJAN	270
PAN	1,435	MORGAN	2,824	MARTIN	223	MIRANDA	140	NAIR	560
PENG	530	MORRIS	3,223	MAYER	1,097	MOLINA	129	NARASIMHAN	225
SHEN	1,480	MURPHY	3,609	MEYER	3,004	MORALES	146	NARAYAN	312
SHI	964	MURRAY	2,207	MOLNAR	335	MORENO	128	NARAYANAN	419
SHIH	938	MYERS	2,625	MORIN	320	MUNOZ	177	NATARAJAN	301
SONG	636	NELSON	6,444	MUELLER	2,242	NUNEZ	207	PAREKH	301
SU	1,025	OLSON	3,140	MULLER	985	ORTEGA	206	PARIKH	286
SUN	2,521	PARKER	3,181	NAGEL	383	ORTIZ	362	PATEL	3,879
TAI	463	PETERSON	4,912	NATHAN	171	PADILLA	116	PATIL	352
TAM	589	PHILLIPS	3,875	NILSSEN	234	PAZ DE ARAUJO	148	PRAKASH	326
TAN	1,105	PRICE	2,062	NOVAK	788	PEREIRA	280	PRASAD	549
TANG	2,277	REED	2,645	PAGANO	177	PEREZ	675	PURI	233
TENG	437	RICHARDSON	2,114	PALERMO	177	QUINTANA	126	RAGHAVAN	378
TONG	677	ROBERTS	4,352	PASTOR	238	RAMIREZ	345	RAHMAN	367
TSAI	1,244	ROBINSON	3,741	POPP	202	RAMOS	226	RAJAGOPALAN	396
TSANG	499	ROGERS	2,974	RAO	343	REGNIER	137	RAMACHANDRAN	388
TSENG	538	ROSS	2,377	REITZ	248	REIS	168	RAMAKRISHNAN	270
TUNG	565	RUSSELL	2,611	ROHRBACH	246	REYES	150	RAMAN	222
WANG	11,905	RYAN	2,404	ROMAN	362	RIVERA	489	RAMASWAMY	244
WEI	1,317	SCOTT	3,583	ROSTOKER	245	RODRIGUES	188	RAMESH	364
WEN	455	SHAW	2,369	SCHMIDT	3,753	RODRIGUEZ	1,314	RANGARAJAN	244
WONG	4,811	SIMPSON	2,014	SCHNEIDER	2,246	ROMERO	292	RAO	1,196
WOO	710	SMITH	24,173	SCHULTZ	2,273	RUIZ	297	REDDY	459
WU	5,521	SNYDER	2,335	SCHULZ	921	SALAZAR	179	ROY	279
XIE	609	STEVENS	2,221	SCHWARTZ	2,394	SANCHEZ	717	SANDHU	878
XU	2,249	STEWART	2,924	SCHWARZ	633	SANTIAGO	158	SAXENA	213
YAN	826	SULLIVAN	2,933	SPERANZA	215	SERRANO	172	SHAH	2,467
YANG	4,584	TAYLOR	6,659	SPIEGEL	177	SILVA	457	SHARMA	1,249
YAO	699	THOMAS	5,312	STRAETER	454	SOTO	158	SINGH	2,412
YE	525	THOMPSON	6,424	THEEUWES	247	SOUZA	145	SINGHAL	245
YEE	729	TURNER	2,855	TROKHAN	167	SUAREZ	150	SINHA	463
YEH	928	WALKER	4,887	VOCK	423	TORRES	352	SIRCAR	225
YEN	467	WALLACE	1,963	WACHTER	199	VALDEZ	127	SRINIVASAN	876
YIN	617	WARD	2,913	WAGNER	2,499	VARGA	130	SRIVASTAVA	498
YU	2,293	WATSON	2,139	WEBER	3,003	VASQUEZ	153	SUBRAMANIAN	702
YUAN	825	WHITE	6,190	WEDER	1,067	VAZQUEZ	260	THAKUR	381
ZHANG	4,532	WILLIAMS	10,442	WEISS	1,533	VELAZQUEZ	134	TRIVEDI	383
ZHAO	1,337	WILSON	7,677	WOLF	1,604	VINALS	220	VENKATESAN	281
ZHENG	1,037	WOOD	4,525	WRISTERS	185	YU	140	VERMA	262
ZHOU	1,517	WRIGHT	4,521	ZIMMERMAN	1,542	ZAMORA	120	VISWANATHAN	218
ZHU	1,749	YOUNG	5,957	ZIMMERMANN	226	ZUNIGA	128	VORA	223

Table A4: Most Common US Ethnic Surnames (continued)

Japanese		Korean		Russian		Vietnamese	
AOKI	141	AHN	610	AGHAJANIAN	77	ABOU-GHARBIA	22
AOYAMA	66	BAE	122	ALPEROVICH	64	BAHN	15
ASATO	73	BAEK	77	ALTSHULER	71	BANH	21
CHEN	88	BAK	68	ANDREEV	94	BI	158
DOI	90	BANG	91	ANSCHER	95	BICH	18
FUJII	92	BARK	39	BABICH	79	BIEN	91
FUJIMOTO	98	BYUN	87	BABLER	73	BUI	309
FUKUDA	84	CHA	45	BARINAGA	72	CAN	19
FURUKAWA	218	CHAE	33	BARNA	96	CONG	41
HANAWA	69	CHANG	289	BELOPOLSKY	71	DANG	23
HARADA	90	CHIN	33	BERCHENKO	94	DIEM	24
HASEGAWA	171	CHO	977	BLASKO	79	DIEP	52
HASHIMOTO	110	CHOE	193	BLONDER	82	DINH	232
HAYASHI	148	CHOI	1,081	BONIN	97	DIP	11
HEY	75	CHON	33	CODILIAN	90	DO	13
HIGASHI	98	CHOO	94	COMISKEY	74	DOAN	616
HIGUCHI	81	CHUN	330	DAMADIAN	118	DOMINH	33
HONDA	102	CHUNG	1,499	DANKO	69	DONLAN	21
IDE	136	DROZD	45	DAYAN	143	DOVAN	26
IKEDA	98	EYUBOGLU	36	DERDERIAN	169	DUAN	241
IMAI	129	GANG	34	DOMBROSKI	66	DUE	20
INOUE	90	GU	533	ELKO	81	DUONG	153
IRICK	86	HAHM	42	FETCENKO	62	DUONG-VAN	13
ISHIDA	93	HAHN	1,016	FISHKIN	82	ESKEW	12
ISHII	82	HAM	45	FOMENKOV	73	GRAN	20
ISHIKAWA	208	HAN	145	FRENKEL	71	HAC	20
ITO	260	HANSELL	39	FRIDMAN	67	HAUGAN	16
IWAMOTO	78	HOGLE	43	FROLOV	68	HO	35
KANEKO	157	HONE	78	GARABEDIAN	104	HOANG	277
KATO	113	HONG	907	GELFAND	139	HOPPING	15
KAUTZ	87	HOSKING	63	GINZBURG	73	HUYNH	317
KAWAMURA	87	HUH	32	GITLIN	73	HUYNH-BA	19
KAWASAKI	104	HWANG	108	GLUSCHENKOV	73	KHA	13
KAYA	78	HYUN	54	GORALSKI	69	KHAW	20
KIMURA	108	IM	80	GORDIN	65	KHIEU	35
KINO	74	JANG	46	GORIN	99	KHU	13
KINOSHITA	93	JEON	134	GRINBERG	104	KHUC	15
KIRIHATA	107	JEONG	122	GROCHOWSKI	77	LAHUE	17
KISHI	65	JI	268	GUREVICH	107	LAURSEN	72
KIWALA	132	JIN	673	GURSKY	89	LAVAN	18
KOBAYASHI	296	JO	41	GUZIK	79	LE	1,263
LI	75	JOO	68	HABA	96	LE ROY	29
LIU	84	JU	55	HYNECEK	82	LEEN	75
MAKI	167	JUNG	582	IBRAHIM	229	LEMINH	17
MATSUMOTO	147	KANG	809	IVANOV	165	LUONG	107
MIYANO	70	KIANI	74	IVERS	66	LY	118
MIZUHARA	87	KIM	5,455	JOVANOVIC	65	MINH	41
MORI	128	KO	595	JU	126	NELLUMS	17
MORITA	64	KOO	214	JUHASZ	71	NGO	735
MOSLEHI	165	KUN	63	KAHLE	173	NGUY	12
MOTOYAMA	130	KWAK	96	KAMINSKI	393	NGUYEN	4,720
MURAKAMI	67	KWON	298	KAMINSKY	150	NHO	12

Table A4: Most Common US Ethnic Surnames (continued)

Japanese		Korean		Russian		Vietnamese	
NAJJAR	81	LEE	1,032	KANEVSKY	114	NIEH	69
NAKAGAWA	125	LIM	135	KAPLINSKY	69	NIM	14
NAKAJIMA	99	MENNIE	96	KAPOSI	72	PHAM	901
NAKAMURA	187	MIN	242	KHAN	104	PHAN	27
NAKANISHI	64	NA	34	KHANDROS	161	PHANG	11
NAKANO	104	NAM	68	KHOVAYLO	69	PHY	19
NEMOTO	70	NEVINS	42	KOLMANOVSKY	70	POSTMAN	12
NISHIBORI	88	NYCE	56	KORSUNSKY	153	QUACH	95
NISHIMURA	131	OH	461	KOWAL	74	QUI	11
NODA	107	PAEK	41	LAPIDUS	63	QUY	13
OGAWA	74	PAIK	144	LEE	113	ROCH	26
OGURA	209	PAK	116	LOPATA	113	TA	91
OHARA	269	PARK	2,145	MESSING	74	TAKACH	30
OHKAWA	89	QUAY	107	METLITSKY	95	TAU	23
OKADA	87	RHEE	191	MIKHAIL	115	THACH	33
OKAMOTO	103	RIM	57	MIRKIN	66	THAI	86
ONO	148	RYANG	38	MOGHADAM	72	THAO	21
OVSHINSKY	314	RYU	99	NADELSON	65	THI	13
SAITO	136	SAHM	45	NAZARIAN	75	THIEN	15
SAKAI	79	SAHOO	58	NEMIROVSKY	73	THUT	28
SASAKI	209	SEO	47	NIE	72	TIEDT	14
SATO	231	SHIM	162	OGG	125	TIEP	12
SETO	73	SHIN	399	PAPADOPOULOS	132	TIETJEN	59
SHIMIZU	103	SHINN	96	PAPATHOMAS	67	TO	76
SUZUKI	306	SIN	62	PETROV	102	TON-THAT	16
TAKAHASHI	245	SJOSTROM	39	PINARBASI	131	TRAN	2,050
TAKEUCHI	242	SO	332	PINCHUK	123	TRANDAI	14
TAMURA	83	SOHN	78	POPOV	81	TRANG	34
TANAKA	328	SON	147	PROKOP	86	TRANK	11
THOR	66	SONG	105	RABER	78	TRIEU	49
TSUJI	92	SUE	64	RABINOVICH	123	TRONG	12
TSUKAMOTO	89	SUH	311	ROBICHAUX	65	TRUC	27
UCHIDA	72	SUK	75	RUBSAMEN	69	TU	545
UEDA	72	SUNG	41	SAHATJIAN	66	TUTEN	23
WADA	153	SUR	38	SARKISIAN	65	TUY	16
WANG	81	TOOHEY	33	SARRAF	82	TY	27
WATANABE	416	UM	36	SCHREIER	62	VAN	58
WU	67	WHANG	175	SCHWAN	81	VAN CLEVE	40
YAMADA	180	WON	108	SIMKO	77	VAN DAM	20
YAMAGUCHI	102	YI	237	SMETANA	69	VAN LE	17
YAMAMOTO	432	YIM	145	SOFRANKO	66	VAN NGUYEN	29
YAMASAKI	67	YOHN	32	SOKOLOV	91	VAN PHAN	26
YAMASHITA	105	YOO	290	SORKIN	111	VAN TRAN	15
YAMAZAKI	91	YOON	614	TABAK	85	VIET	11
YANG	65	YOUN	38	TEPMAN	80	VO	269
YASUDA	75	YU	198	TERZIAN	87	VO-DINH	32
YOSHIDA	178	YUH	96	VASHCHENKO	96	VOVAN	20
YUAN	112	YUM	78	WASILEWSKI	80	VU	502
ZHAO	81	YUN	222	ZEMEL	126	VUONG	107

Table A5: Ethnic Inventors and MSA Characteristics Including Overall Ethnic Shares

	English		Chinese		Indian		European		Hispanic	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent Variable is Share of 1985-2004 Ethnic Patenting in the MSA									
1980 Share of Ethnic	0.336	0.464	1.126	1.137	0.373	0.498	0.324	0.390	0.105	-0.042
Population in MSA	(0.350)	(0.188)	(0.375)	(0.336)	(0.220)	(0.124)	(0.140)	(0.066)	(0.144)	(0.213)
Log Population	0.473	0.162	-0.374	-0.692	0.315	0.003	0.540	0.266	0.790	0.860
of MSA	(0.380)	(0.167)	(0.297)	(0.363)	(0.270)	(0.196)	(0.226)	(0.102)	(0.192)	(0.254)
Log Population	0.040	0.108	0.108	0.366	0.016	-0.057	0.041	0.193	-0.024	-0.097
Density of MSA	(0.028)	(0.099)	(0.052)	(0.210)	(0.047)	(0.185)	(0.039)	(0.094)	(0.033)	(0.151)
Coastal Access	-0.018	0.098	-0.036	-0.144	0.023	0.335	-0.002	0.105	0.048	0.334
of MSA	(0.035)	(0.131)	(0.023)	(0.115)	(0.054)	(0.251)	(0.033)	(0.126)	(0.043)	(0.194)
Share of Population	0.141	0.372	0.082	0.121	0.129	0.428	0.111	0.376	0.089	0.263
with Bachelors Ed.	(0.042)	(0.241)	(0.028)	(0.141)	(0.058)	(0.347)	(0.039)	(0.235)	(0.028)	(0.184)
Share of Population	-0.138	-0.650	-0.142	-0.518	-0.156	-1.132	-0.110	-0.641	-0.060	-0.509
under 30 in Age	(0.072)	(0.537)	(0.066)	(0.247)	(0.116)	(0.913)	(0.068)	(0.555)	(0.045)	(0.430)
Share of Population	-0.086	-0.399	-0.110	-0.339	-0.100	-0.693	-0.051	-0.318	-0.016	-0.251
over 60 in Age	(0.057)	(0.344)	(0.061)	(0.203)	(0.090)	(0.594)	(0.054)	(0.352)	(0.039)	(0.284)
Share of Population	-0.062	-0.386	-0.038	-0.709	-0.058	-0.328	-0.039	-0.238	-0.038	-0.118
Female	(0.026)	(0.265)	(0.026)	(0.333)	(0.039)	(0.472)	(0.023)	(0.236)	(0.021)	(0.199)
Weights	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-Squared	0.79	0.84	0.82	0.84	0.56	0.66	0.81	0.88	0.83	0.90

Notes: See Tables 3 and 4. Estimations incorporate the overall share of each ethnicity in MSAs from the 1990 Census.