How Much Does Immigration Boost Innovation? *

Jennifer Hunt †
McGill University and NBER

jennifer.hunt@mcgill.ca

May 14, 2008

*I am grateful to David Munroe for excellent research assistance, and for helpful comments to Francisco Alvarez–Cuadrado, Leah Brooks, David Card, Lee Fleming, Rachel Friedberg, David Green, Francisco Gonzales, Judy Hellerstein, Chad Jones, Daniel Parent, Giovanni Peri, Steve Pischke, Regina Riphahn, Eric Stuen and Dee Suttiphisal, seminar participants at the NBER and Nürnberg, and several friends holding patents. I thank Jim Hirabayashi of the USPTO and Nicole Fortin for data and Deven Parmar for obtaining and formatting the USPTO data. I am also affiliated with the CReAM, CEPR, IZA and DIW-Berlin, and I acknowledge the Social Science and Humanities Research Council of Canada for financial support.

†Department of Economics, McGill University, 855 Sherbrooke Street West, Montreal, QC, H3A 2T7, Canada; tel. (514) 398-6866; fax (514) 398-4938.

Abstract

I measure the extent to which skilled immigrants increase innovation in the United States by exploring individual patenting behavior as well as state–level determinants of patenting. The 2003 National Survey of College Graduates (NSCG) shows that immigrants patent at double the native rate, and that this is entirely accounted for by their disproportionately holding degrees in science and engineering. These data imply that a one percentage point rise in the share of immigrant college graduates in the population increases patents per capita by 7.7%. This could be an overestimate of immigration's benefit if immigrant inventors crowd out native inventors, or an underestimate if immigrants have positive spill–overs on inventors. Using a 1950–2000 state panel, I show that natives are not crowded out by immigrants, and that immigrants do have positive spill–overs, resulting in an increase in patents per capita of 8–21% in response to a one percentage point increase in immigrant college graduates. I isolate the causal effect by instrumenting the change in the share of skilled immigrants in a state with the initial share of immigrant high school dropouts from Europe, China and India. In both data sets, the positive impacts of immigrant post–college graduates and scientists and engineers are larger than for immigrant college graduates.

Economists have studied the impact of immigration on a variety of host country outcomes. For example, Card (2007) considers U.S. immigration's impact on population growth, skill composition, internal migration, wages, rents, taxes and the ethnic and income composition of neighborhoods and schools. In contrast, the impact of immigration on innovation has received less attention. In addition to the direct contributions of immigrants to research, immigration could boost innovation indirectly through positive spill–overs on fellow researchers, the achievement of critical mass in specialized research areas, and the provision of complementary skills such as management and entrepreneurship. Some tantalizing facts hint at the possible importance of these effects for the United States. Compared to a foreign–born population of 12% in 2000, 26% of U.S.–based Nobel Prize recipients from 1990–2000 were immigrants (Peri 2007), as were 25% of founders of public venture–backed U.S. companies in 1990–2005 (Anderson and Platzer n.d.), and founders of 25% of high–tech companies founded 1995–2005 with more than one million dollars in sales in 2006 (Wadhwa et al. 2007). Immigrants are over–represented amongst members of the National Academy of Sciences and the National Academy of Engineering, amongst authors of highly–cited science and engineering journal articles, and amongst founders of bio–tech companies undergoing IPOs (Stephan and Levin 2001). Kerr (2007) documents the surge in the share of U.S. patents awarded to U.S.–based inventors with Chinese and Indian names to 12% of the total by 2004, and Wadhwa et al. (2007) find that non–U.S. citizens account for 24% of international patent applications from the United States.

The goal of my paper is to assess the impact of skilled immigration on innovation as measured by U.S. patents. The purpose of studying patents is to gain insight into technological progress, a driver of productivity growth and ultimately economic growth. If immigrants increase patents per capita, they are likely to increase output per capita and make natives better off. This is an important consideration for the debate concerning how many and what type of immigrants should be admitted to the United States, and particularly for the discussion of the appropriate number of employer–sponsored H1-B visas for skilled (especially science and engineering) workers. The context of the discussion is the shift from European to low and middle–income source countries since the Immigration Act of 1965, and the concomitant faster increase in unskilled immigration than skilled immigration.

I begin by analyzing how much immigrants patent using the 2003 National Survey of College Graduates (NSCG), which contains information on both patenting activity and birth place. The individual–level data allow me to gauge the impact of immigrants on patents per capita under the assumption that immigrants do not influence the behavior of natives or other immigrants, and allow me to examine whether immigrants patent more than natives because they have higher ability or merely different education. In order to account for immigrants' possible influence on natives or other immigrants, I turn to a panel of U.S. states from 1950–2000, based on data from the U.S. Patent and Trademark Office, the decennial censuses and other sources. I test whether skilled immigrants crowd out skilled natives from the states (and occupations) to which they move, and I provide estimates of skilled immigrants' impact on patents per capita that encompass both immigrants' own patenting and any positive spill–overs immigrants might have. To obtain the causal effect of immigrants despite their endogenous choice of destination state, I difference the data across census years, and instrument the change in the share of skilled immigrants in a state with the state's initial share of immigrant high school dropouts from Europe, China and India, the origin regions of at least 42% of skilled immigrants throughout the period. All of my analysis captures invention at companies, universities and government laboratories, and the contributions of immigrants arriving before or after their tertiary education.

I contribute to three understudied areas: the impact of immigration on innovation, the regional determinants of innovation and the individual determinants of innovation. My work is also relevant for the macroeconomic growth literature, where the link between innovation and the number of researchers is the key to growth.[1] My results also reveal the importance of convergence in innovation across states, surely an important mechanism behind the convergence of per capita personal income analyzed in papers beginning with

[1]Aghion and Howitt (1992), Grossman and Helpman (1991a,b), Jones (1995), Romer (1990).

Barro and Sala-i-Martin (1991).

I go beyond the most closely related paper linking immigration and innovation, Peri (2007), by extending the state panel, using instrumental variables, defining skilled immigration consistently across time and more broadly, testing for crowd–out of natives and adding individual–level analysis. These considerations also distinguish my paper from the time–series analysis of Chellaraj, Maskus and Matt (2004). Both of these papers find skilled immigration increases U.S. patenting. My analysis is more general than that of Stuen, Mobarak and Maskus (2007), who find that immigrant students increase U.S. university patenting and science and engineering publishing. A related paper by Niebuhr (2006) concludes that German regions with more diverse worker nationalities (as measured by the Herfindahl index) patent more. The result is not robust to region fixed effects, however, no doubt in part because she has only two years of data close in time (1997 and 1999).

While a small number of papers studies the regional determinants of patenting, to the best of my knowledge none save Peri (2007) and Niebuhr (2006) exploits either a panel dimension or a period of several decades. The most closely related paper is by Zucker and Darby (2006), who find for 1981–2004 that a Bureau of Economic Analysis (BEA) region's non–university patenting is unaffected by the presence of star scientists, a high wage (which they view as proxying for education) and a high stock of relevant journal publications. Bottazzi and Peri (2003) average over 1977–1995 to obtain cross–section evidence of geographic spill–over effects of R&D spending on patenting in European regions.[2] I am not aware of previous papers with regression analysis of the individual determinants of patenting, though Morgan, Kruytbosch and Kannankutty (2003) note in passing the immigrant advantage in patenting in the 1995 NSCG, and economic historians have studied the characteristics of nineteenth century inventors (e.g. Khan and Sokoloff 1993).[3]

[2]See also descriptive statistics in Hicks et al. (2001).

[3]Other relevant papers include Agrawal, Kapur and McHale (2002), who find that emigration from India reduces access to knowledge in India, Zucker et al. (2006), who examine the determinants of a BEA region's publications in nanotechnology, and Marx, Strumsky and Fleming (2007) and Stuart and

My analysis of the NSCG data shows that immigrants account for 24% of patents, twice their share in the population, and that the skilled immigrant patenting advantage over skilled natives is entirely accounted for by immigrants' disproportionately holding degrees in science and engineering fields. The results imply that a one percentage point increase in skilled immigrants' share of the population increases patents per capita by 7.7%. This could overestimate the contribution of immigrants, if immigrants crowd out natives, but using the panel of states I show this does not happen, at least in the long run. This is consistent with Borjas (2006), who finds that immigrants do not crowd out natives as a whole from graduate school. Instead, the state panel data show evidence of positive spill–overs of natives, since the estimates of the immigrant impact on patents per capita are higher than in the NSCG: a one percentage point rise in the share of immigrant college graduates in the population increases patents per capita by 8–21%. The state–level results mean that the 1990–2000 increase in the population share of this group from 2.2% to 3.5% increased patents per capita by 10–26%. Consistent with the individual–level analysis, I find that immigrants have more than double the impact on innovation that natives do. I find that immigrants who are scientists and engineers or who have post–college education boost patents per capita more than immigrant college graduates.

1 Methodology

I use individual–level data to measure and explain differences in patenting behavior between immigrants and natives, and to gauge the contribution of immigrants to patenting per capita under the assumption that immigrants do not affect the behavior of natives or other immigrants. I then use state–level data to test for crowding out of natives by immigrants, and to estimate the effect of immigrants on patenting per capita, including any positive spill–overs.

Sorenson (2003), who examine the effect of a state's enforcing non–compete laws on inventor inter–firm mobility and biotech IPOs respectively. The main focus of the literature on geography and innovation is on geographic patterns of patent citing (see Jaffe, Trajtenberg and Henderson 1993 and successor papers).

1.1 Individual–level data

A simple measure of the percent increase in patenting per capita owing to skilled immigrants is the percent by which they increase patents minus the percent by which they increase the population. This increase can then be divided by the skilled immigrants' share in the population to yield the percent increase in patents per capita attributable to each percentage point of skilled immigrants in the population.

I show below that the difference in patenting behavior between immigrants and natives comes through the difference in the probability of patenting at all, not in the number of patents conditional on patenting. To explore the reasons for higher immigrant patenting, I therefore estimate a probit for the probability of having a patent granted, or the probability of commercializing or licensing a patent, weighted by the survey weights:

$$P(patent_j) = \alpha + \beta_1 I_j + X_j \beta_2 + \epsilon_j, \tag{1}$$

where j indexes individuals and I is a dummy for the foreign–born. The coefficient of interest is β_1. I am interested in how much of the raw patenting gap between immigrants and natives (the value of β_1 with no X covariates) can be explained by adding the covariates X: field of study of the highest degree, the highest degree, and demographic variables.

1.2 State–level data

I supplement this analysis using a panel of U.S. states with decennial data from 1950–2000. I have chosen the geographic dimension of variation to supplement the time dimension as I do not have an instrument for immigrant share by firm, the usual unit of observation in patent studies. By extending the period of observation back to 1950, I am able to distinguish long run and short run effects by differencing the data in lengths varying from ten to 50 years.[4] I do not extend it to prior decades as patenting in the years of the Great Depression and the Second World War was probably atypical.

[4]Strictly speaking, I should refer to low–frequency and high–frequency effects.

In order to obtain an estimate of the impact of immigrants on innovation that ecompasses both their own inventions and any positive spill–over effects, I estimate

$$\Delta log\ \frac{P_{i,t+1}}{POP_{i,t+1}} = \alpha + \gamma_1 \Delta I^S_{it} + \gamma_2 \Delta N^S_{it} + \Delta X_{it}\gamma_3 + \gamma_4 Z_{i,1950} + \mu_t + \Delta\epsilon_{it}, \qquad (2)$$

where i indexes states, P is the number of patents, POP is state population, I^S is the share of the population or workforce (18–65) composed of skilled immigrants, N^S is the corresponding share for natives, $Z_{i,1950}$ are characteristics of the state in 1950, X are contemporaneous state characteristics and μ_t are year dummies. The coefficient of interest is γ_1, though its size relative to γ_2 is also of interest. I also present results from specifications where the dependent variable is not in logs.

I define a skilled person variously as one with a college degree or more, one with post–college education, or one working in a science, engineering or computer science occupation. I include characteristics of the state in 1950 (including land area), as the other covariates do not appear to capture the convergence in patents per capita occurring over the time period. The X covariates comprise the log of defense procurement spending and the log of the average age of state residents (18–65). I deliberately do not include total R&D spending (including companies' spending), as I believe this to instead be a potential outcome variable. I lead the dependent variable by one year to allow for a year of research time between the change in the inputs and the patent application, as anecdotal evidence suggests the lag can vary between a few months and two years.

There were several major changes to the patent system between 1980 and 1998 (see Hall 2005). One change led to a large increase in patenting in electrical engineering relative to other fields. To capture potentially differential effects of this by state, I include among the X's the share of employment in electrical engineering–related fields in 1980, interacted with year dummies.[5] I use state populations to weight the regressions,[6] since in some small states one company drives the time series of patenting,[7] and I cluster standard errors by

[5] I use 1980 values as electrical engineering employment was still tiny in most states in 1950–1970.

[6] Specifically, I weight by $1/(1/pop_{i,t+1} + 1/pop_{i,t-k+1})$, where k is the length of the difference.

[7] Idaho's emergence as the state with most patents per capita has been driven by one semi–conductor company, Micron Technology Inc., founded in 1978, which was granted 1643 patents in 2001 and was the

state to allow for serial correlation.

Because I account for state fixed effects by estimating equations differenced across time, I elect not to include the change in the patent stock among the regressors as would be predicted by patent models. Furthermore, because I analyze long–run changes, I have chosen not to use a partial adjustment model.[8]

Equation 2 suffers from an endogeneity problem. Skilled workers are likely to migrate to states which are growing or innovating, causing $\hat{\gamma}_1$ and $\hat{\gamma}_2$ to be biased up in least squares estimation. On the other hand, $\hat{\gamma}_1$ in particular could be biased towards zero owing to measurement error.[9] I use several sets of instruments to address these problems for skilled immigrants. To instrument $\Delta I^S = I^S_t - I^S_{t-k}$, I use I^{HSD}_{t-k}, the share of the population that is a high school dropout at time $t-k$, and its square. The presence of immigrant high school dropouts in a state will mean the existence of cultural amenities attractive also to skilled immigrants. On the other hand, high school dropouts should play a minimal role in innovation, justifying their exclusion from equation 2. A (preferred) variant of this instrument set is three variables for the share of high school dropouts at $t-k$ who were born in Europe, China and India, the most common source regions for skilled immigrants. Alternatively, I use the values of the variables at time $t-k-10$ as instruments so as to be more confident that they are unaffected by unobserved factors influencing the change in patenting between $t-k$ and t.[10]

I also use the state panel to test for crowd–out of natives, which if present would affect the impacts calculated using both the individual–level and state–level data. Natives may choose not to enter careers in science and engineering, or to work less, owing to

fourth–ranked company in this regard.

[8]I have estimated these models. The coefficient on the change in the stock of patents is close to one, rendering all other coefficients insignificant, while the coefficient on the partial adjustment term is insignificant.

[9]There is considerable measurement error for small states in the 1950 census, which was a smaller sample than later years and which asked certain key questions of only one quarter of the sample. There may also be measurement error for the share of immigrant scientists and engineers in all years.

[10]For ΔN^S (native skilled workers), I have experimented unsuccessfully with lagged college enrollments as an instrument. The enrollment data only begin in the 1970s in any case.

competition from immigrants whose comparative advantage is in less language–intensive and less institution–specific occupations. Any drop in native inventors must be taken into account when calculating the net benefit of immigrants. A more complex concern is that native innovators forego migration to states with many immigrant inventors. If by doing so native inventors forego an increase in productivity, this too must be taken into account when calculating the net benefit of immigrants. For this to be logical, natives must be foregoing migration because of personal distaste for foreigners, since if they are foregoing a productivity increase they must also be foregoing a wage increase. Natives without such a distaste will be attracted to states with skilled immigrants if the skilled immigrants convey positive productivity spill–overs.

I test for both types of crowd–out using the simple approach of Card (2005) by running the regression

$$\Delta S_{it} = a + b\Delta I^S_{it} + f\Delta Age_{it} + \mu_t + \Delta\nu_{it}, \tag{3}$$

where S is the share of the population or workforce (aged 18–65) composed of skilled natives and immigrants, I^S is again the share of skilled immigrants, and Age is the average age of the state's population between 18 and 65. I control for the average age of the state since birth cohort is the strongest determinant of schooling. If increases in the skilled immigrant share translate into one for one increases in the total skilled share, there is no crowd–out and $\hat{b} = 1$. Complete crowd–out would be represented by $\hat{b} = 0$, while $\hat{b} > 1$ would indicate that skilled natives were attracted to states with many skilled immigrants. Measurement error could also cause $\hat{b}$ to be less than one.

2 Data and Descriptive Statistics

2.1 Individual–level data

I use the individual–level data from the 2003 National Survey of College Graduates (NSCG). These data are a stratified random sample of people reporting having a bachelor's degree or higher on the long form of the 2000 census. In 2003, all respondents who had ever worked were asked whether they had applied for a U.S. patent since October 1998, whether they had been granted any U.S. patent since October 1998, and if so, how many, and how many had been commercialized or licensed.[11] The survey will not capture patents by those with less than a college degree, but I assume that most patents are captured. The Data Appendix provides more information on the NSCG. I include in my sample respondents 65 or younger (the youngest respondent is 23, but few are younger than 26). I define immigrants as the foreign born.

I define three (not mutually exclusive) skill categories, motivated in part by consistency with categories that can be distinguished in the censuses: college graduates (i.e. the full sample); holders of a post–college degree; and those working as scientists and engineers in the survey week. Only 56% of respondents who had been granted a patent reported working in a science or engineering occupation. Another 20% reported a management occupation: a research team's manager is sometimes listed as a co–inventor on a patent, and all inventors listed are captured in the data, and many inventors will have been promoted to management since obtaining a patent. Science and engineering technicians represent 2.8% of patent holders, and respondents in health–related occupations represent another 3.3%.

Table 1 shows details of how patenting varies by immigrant status for the three skill groups. For college graduates (the whole sample, columns 1–2), 1.9% of immigrants were granted patents, compared to 0.9% of natives, and patents per capita were 0.058 for immigrants and 0.028 for natives, an immigrant/native ratio of 2.1 in both cases. Immigrants held a similar advantage in patents commercialized or licensed, patents likely to benefit society more than others: 1.2% immigrants had commercialized a patent compared to 0.6% for natives, and commercialized patents per capita were 0.029 for immigrants and 0.017 for natives. Immigrants therefore patent at about twice the native rate, with the difference being in the probability of patenting at all. The immigrant–native gap is larger

[11]Questions on patents were also asked in the 1995 NSCG, but only of respondents who said they worked in research and development in the survey week, which will cause the patents of job changers to be missed.

for the sample with post–college education (columns 3–4), but much smaller for the sample working in science and engineering occupations (columns 5–6). For example, 6.3% immigrants in the latter sample had been granted a patent, compared to 4.9% natives, and immigrants hold 1.37 times the patents per capita of natives. Appendix Table 1 contains the means of variables used in the regression analysis below.

2.2 State–level data

The patent data used in the state–level analysis come from the U.S. Patent and Trademark Office (USPTO). Patents are attributed to states based on the home address of the first inventor on the patent. I merge a series based on electronic data from 1963 onwards with a series from paper records for 1883–1976. The two series are not completely comparable, and the merging details are given in the Data Appendix. Patents are classified according to application (filing) date. Figure 1 shows the evolution of total patents and patents per 100,000 residents from 1951-2001. The patent decline of the 1970s was caused by bottlenecks at the USPTO (Griliches 1990), while the more recent growth has been driven by growth in patents related to electrical engineering (Hall 2005).

In Figure 2 I use patent data from 1929 to 2001 to display the long–run convergence across states in patenting, as measured by changes in the (unweighted) standard deviation of log patents. The convergence in patents, shown by the downward slope of the top line, is not merely a function of convergence in population, as is demonstrated by the convergence in patents per capita (bottom line). However, there is divergence in patents per capita from 1990–2001, and there have historically been other periods of divergence. California is a force for divergence, as may be seen by the growing gap between the inequality of state patent counts (top line) and the inequality of counts without California (middle line).

I have also used the NBER Patent Citation Data File (Hall, Jaffe and Trajtenberg 2001), as updated by Hall, which contains the fields of patents awarded 1963–2002 and citations to them since 1975. These data permit patents to be weighted by citations to yield a quality–adjusted patent measure by application year. The data are imperfect for my decennial analysis requiring patents counts for 1950–2000. I have performed analysis with the independent variables for 1970–2000 and patent counts for 1970, 1980, 1990 and 1997, since 1997 is the latest application year for which most patents had been awarded by 2002. 1970 is also imperfect, as most citations to patents applied for in that year are not recorded.

To compute the shares of the population in various education and occupation classes, to divide these into immigrant and native, to calculate the average age of the state's population and to obtain weekly wages, I use the IPUMS microdata of the decennial censuses. I base most calculations on the population or workforce aged 18–65. Post–college education is the highest education level that can be measured consistently throughout 1950–2000. I define immigrants to be the foreign born. Information for Alaska and Hawaii is not available in 1950.

I also use the state population and state personal income per capita (state gross product is not available for the whole period) from the Bureau of Economic Analysis (BEA). I use Department of Defense data on the value of Prime Military Contracts (defense procurement contracts) from 1951 to the present, and attribute the 1951 values to 1950.

Full details on the data construction are given in the Data Appendix, while the variable means, weighted by population, are reported in Table 2. Between 1950 and 2000, the share of the population 18–65 composed of immigrants with college education or more increased tenfold to 3.5%, while the equivalent share for post–college increased eightfold to 1.6%. The population shares comprising natives with at least college and with post–college increased from 6.2% to 20.0% and from 2.3% to 7.7% respectively. The share of workers composed of immigrant scientists and engineers multiplied ninefold to 0.9%, while the native share rose from 1.2% to 3.5%. The Appendix Table 2 contains the means of the variables used as instruments.

3 Results

3.1 Individual determinants of patenting

The NSCG data may be used to estimate the direct effect of immigration on patenting, ignoring possible crowd–out or spill–over effects. Immigrants hold 24% of patents in the (weighted) data, and in the 2000 census (the basis of the NSCG sampling frame), college–graduate immigrants were 3.5% of the U.S. population. Skilled immigrants thus increase patents by 31.5% and the population by 3.6%, yielding an increase in patents per capita of 27%, or 7.7% per percentage point of skilled immigrants in the population.

As immigrants with post–college education have 1.9 (=0.113/0.058) times as many patents per capita as immigrants with only a college degree (see Table 1), the direct impact of an extra percentage point of post–college immigrants in the population is likely to be about 1.9 times higher, or an extra $1.9 \times 7.7 = 15\%$. Similarly, the contribution of an additional percentage point scientists and engineers is likely to be $3.1 \times 7.7 = 24\%$.

It is useful to know whether immigrants patent more because of higher innate ability or different observable characteristics. The fact that in Table 1 immigrants' patenting advantage over natives is much smaller in the scientist and engineer sample (columns 5 and 6) than in the overall sample (columns 1 and 2) suggests that immigrants' advantage is due in large part to a greater science and engineering orientation. Table 3 lends further support to this. Column 1 shows that, for the whole sample, 6.6% of those with a highest degree in physical science and 6.0% of those with a highest degree in engineering had patented, far ahead of other fields. Column 2 shows a qualitatively similar picture for commercialized or licensed patents. Columns 3 and 4 show that the share of immigrants with physical science and engineering degrees is more than twice as high as for natives.

In Table 4, I pursue this explanation with the aid of a probit for the probability of patenting. Column 1 shows that immigrants are 1.0 percentage points more likely to have been granted a patent in the sample of college graduates (top panel), 2.3 percentage points more likely in the sample of post–college graduates (second panel) and 1.4 percentage points more likely in the sample of scientists and engineers (third panel). In the second column, I control for 30 dummies for the field of study of the highest degree obtained by the respondent. For college and post–college graduates, the immigrant advantage is reduced to a tenth or less of its original magnitude (though the gap is still statistically significant). For scientists and engineers, who have less variation in their field of study, the gap is reduced by about two thirds (and the remaining gap is statistically insignificant). In the third column, I control for the highest degree obtained by the respondent. For college and post–college graduates, this reduces the already small immigrant–native gap to insignificance. For scientists and engineers, the gap is reversed: immigrants are a statistically significant 0.9 percentage points less likely to patent than natives. Controlling for age, age squared, sex and current employment status in column 4 changes little. Skilled immigrants' advantage is therefore entirely due to the nature of their education, and not to any selection on unobservables such as ability.[12] In columns 5 and 6 I show that the same conclusions may be drawn for the probability of commercializing or licensing a patent.

3.2 Crowd–out

To test for crowd–out, I estimate equation 3. The results with college or more as an indicator of skill are reported in Panel A of Table 5. Column 1 shows that with weighted least squares and ten–year differences, a one percentage point increase in the share of the population that is immigrant college graduates only increases the overall share of college graduates by 0.51 percentage points. This indicates crowd–out, though the coefficient is not statistically significantly different from one. As I increase the length of the differences, evidence of crowd–out disappears: the coefficient is 0.75 for 30–year differences in column 2, and 0.95 for 50–year differences in column 3. In columns 4–6, I report the corresponding instrumental variables results, using the shares of European, Chinese and Indian dropouts at $t-k$ as instruments. The coefficients are smaller – and in the case of ten–year differences significantly different from one – but also increase as the difference length increases, to 0.79 in column 6.

[12]It is possible that unobservable effects cancel out e.g. immigrants may have higher ability but lower quality education.

In panel B I repeat the regressions using the share of post–college educated in the population as the measure of skill. The weighted least squares coefficients in columns 1–3 are significantly greater than one, suggesting that skilled natives are attracted to states (or education levels) with many immigrants ("crowd–in"). With instrumental variables, the point estimate suggests crowd–out for first differences (though the coefficient of 0.63 in column 1 is not statistically significantly different from one), but crowd–in at longer differences.

In panel C I repeat the regressions using the share of workers who are scientists and engineers. The weighted least squares coefficients indicate no (ten and 50–year differences) or little (30 year) crowd–out in columns 1–3. The instrumental variables coefficients in columns 3–6 indicate statistically significant crowd–out for ten and 30–year differences, but by 50–year differences the coefficient is 0.76 and insignificantly different from one.

I have repeated all the regressions including dummies for seven BEA regions and the results change little. In summary, it appears there is some crowd–out of skilled natives by skilled immigrants in the short run but not the long run, and that for post–college graduates the opposite occurs: skilled natives are attracted to skilled immigrants. Though I cannot correct for crowd–out when analyzing the determinants of innovation, these results mean that the short–run benefits of skilled immigrants may be overestimated.

3.3 State determinants of patenting

The evolution of a state's patents over the period 1950–2000 is strongly related to conditions in 1950, as might be expected given the convergence depicted in Figure 2. This is illustrated by Figure 3, which measures the difference in log patents per capita over the period on the y–axis. The x–axis is the log population density of the state in 1950 (population divided by land area in square kilometers). The 1950 population density explains 31% of the weighted variance in 1950–2000 patent growth, and the regression line is downward sloping: states which were densely populated in 1950, which is presumably propitious for innovation, had lower patent growth than lightly populated states. This motivates my inclusion in the regressions below of state land area, 1950 population and 1950 state personal income per capita.

In Table 6, I estimate the state determinants of patenting using differences of different lengths, with a college degree as the measure of skill. In columns 1–4 the dependent variable is the log of patents per capita. The coefficients on the share of immigrant college graduates are positive and significant. In columns 1–3, where I use weighted least squares, a one percentage point increase in the share of the population composed of immigrant college graduates is associated with an 11–12% increase in patenting for ten and 30 year differences, and a 15.6 log point (14%) increase for 50 year differences. These effects are larger than the 7.7% impact calculated based on the NSCG data, implying positive spill–over effects of immigrants.

In column 4, I present the results of instrumenting the ten–year change in skilled immigrant share with the share of European, Chinese and Indian high–school dropouts at $t-10$ (the initial year of the pair of years differenced). The coefficient on the change in the immigrant share is a statistically significant 17.7, and larger than its least squares counterpart of 11.4 in column 1 (though not statistically significantly so). This may indicate that in least squares, measurement error's bias towards zero is more important than upward bias due to the endogenous location choice of immigrants, a possibility mooted by Card and DiNardo (2000) in a similar context. Another possibility is that the instrumental variables estimators place more weight on later years of the sample when the effects seem to be higher (the shares of Chinese and Indian immigrants are more significant in the first stage than the share of Europeans). It could also be that skilled immigrants whose behavior is affected by the instrument (skilled immigrants whose location decision is affected by the presence of other immigrants) are more inventive than other immigrants.

I do not present instrumental variables estimates for longer differences for this or later specifications: results are similar for 20–year differences, whereas for 30–50 year differences the instruments are not strong in the first stage. However, the crowd–out results of Table 5 suggested that to avoid overstating benefits of immigrants, results from longer differences should be preferred. I am obliged to assume that since the short and long–run correlations are similar, the short and long–run causal effects are similar.

In columns 5–7 the dependent variable is simply the change in patents per capita (in these columns the coefficients are multiplied by 100). The least squares effects for ten and 50–year differences are similar: a one percentage point increase in the skilled immigrant share is associated with a 0.000039 increase in patents per capita, which is a 17% increase compared to the mean. The corresponding impact for the unreported 30–year differences is 13%, so the results are similar to those of the log specification in columns 1–3. The instrumental variables estimator in column 7 is larger than its least squares counterpart in column 5, but insignificant. The skilled immigrant coefficients in columns 5–7 are not very sensitive to the covariates included, while the results in columns 1–4 are much smaller if the 1950 covariates (and land area) are not included.

By contrast, most of the coefficients on the change in the share of native college graduates are small and insignificant. The point estimate increases as the difference length increases, and for 50–year differences the coefficient is a significant 6.7 in column 3. As the share of native college graduates changes only gradually (i.e. at low frequency), the absence of significance at short differences probably reflects the emphasis of short differences on high–frequency events (Baker, Benjamin and Stanger 1999). The ratio of the immigrant to the native effect for 50–year differences in columns 3 and 6 is similar to the ratio of patents per capita in the NSCG data: 2.2–2.3 compared with 2.1.

Older populations appear to be more innovative, as indicated by the positive coefficients on the average age of the state in the log specifications of columns 1–4. This may reflect the importance of management or other skills complementary to innovation. As suggested by time series work in Griliches (1990), Department of Defense spending lowers patenting in the log specifications, presumably in part because military invention is primarily protected by secrecy rather than patents. Finally, the importance of the 1950 conditions (and land area) increases with the difference length.

These regressions are repeated in Table 7 with post–college education (panel A) and a science and engineering occupation (panel B) as measures of skill. The least squares coefficients for immigrant post–college range from 17–27 in columns 1–3, where the dependent variable is in logs. These estimates are almost twice as high as for college graduates in Table 6, consistent with the NSCG data. The ten–year difference instrumental variables coefficient in column 4 is higher, at 38.1, but insignificant. The coefficients in columns 5–6 are insignificant, but are also about double their counterparts in Table 6. The instrumental variables coefficient in column 7 is larger than its least squares counterpart in column 5, but also insignificant. The coefficients on the share of native post–college educated are never statistically significant, though the point estimates are higher for the longer differences. The immigrant/native ratio at 50–year differences is 2.8–3.3, compared to 3.1 in the NSCG.

In panel B, the coefficients are significant in all columns for immigrants and most columns for natives, and are larger than for the other skill groups. For immigrants in columns 1–3, a one percentage point increase raises patents by 48–59 log points, or 40–46%. Unlike for college graduates and post–college educated, the instrumental variables estimates (columns 4 and 7) are fairly similar to the least squares estimates (columns 1 and 5). The coefficients are high compared with the direct NSCG effect of about 24% and compared with that of natives at 50–year differences (29 log points or 26%), given that in the NSCG the immigrant patenting advantage over natives was only 37% (which would imply an immigrant coefficient of only 40 in Table 7). The state–based estimated impact of immigrant scientists and engineers should thus be viewed with caution. Various unreported specification checks for 50–year differences do not resolve the discrepancy.

Focusing on least squares estimation, I find similar results to Tables 6 and 7 when I use citation–weighted patent counts for the 1970–2000 period, and I find that the impact of skilled immigrants is largest in the fields of electrical engineering and computer science. These unreported results should be viewed as preliminary while awaiting the updating of the NBER patent data file, but indicate that immigrants contribute as much to high–quality patents as to low–quality patents, consistent with the individual regressions for commercialized and licensed patents.[13]

[13] In other unreported regressions focusing on 1970–2000, I have used the value of federal R&D spending, which is only available by state from 1968 to the present. Its coefficient was insignificant and the variable's

In Table 8 I present various plausible alternative estimates of the effects of skilled immigrants, concentrating on the college–educated and the scientists and engineers, on ten-year differences, and on the log specification (results without logs display similar patterns). I report only the coefficient on the change in the skilled immigrant share, each from a different regression. In the first row, I reproduce the baseline least squares and instrumental variables results from Tables 6 and 7. In the next three rows I vary the instruments used. In the second row I use as instruments the shares of the population composed of European, Chinese and Indian high school dropouts at $t-20$, instead of at $t-10$ as in the baseline. The resulting coefficients are slightly larger than the baseline instrumental variables coefficients, which in turn were higher than the least squares results. In rows 3 and 4, I use as instruments the share of all foreign–born high school dropouts and its square, at $t-10$, and at $t-20$. With these instruments, the point estimates are quite similar to the baseline least squares results for college graduates, but for scientists and engineers are slightly higher than the baseline instrumental variables results.

In the next two rows I experiment with adding covariates. In row 5 I allow for (seven) BEA–region specific trends in patents per capita. This reduces the coefficients to 68–84% of the magnitudes of the baseline row and renders them statistically insignificant, though the least squares coefficients are significant at the 10% level. In row 6, I add instead the interactions of the 1980 share of employment in electrical engineering–related sectors interacted with year dummies. This yields estimates that are also lower than those in the baseline row, though generally statistically significant, this time 79–86% of the baseline magnitudes.

In row 7 I investigate the influence of California in the baseline specification by dropping that state. This reduces the estimates greatly. Finally, I assess the robustness to dropping the 1990–2000 differences (while retaining California), using the baseline specification. This causes the weighted least squares coefficients to become much smaller and insignificant, with point estimates of 6.5 and 21.2 in columns 1 and 3. However, for the college educated, the larger instrumental variables estimate of 12.5 is statistically signifi-

inclusion did not affect other coefficients.

cant in column 2. Instrumental variables point estimates are also much larger than least squares estimates for scientists and engineers. The sensitivity to the dropping of the year 2000 is present at all lengths of differences (these results are not reported). The coefficient on the change in the share of skilled natives, by contrast, is not greatly affected by the dropping of the year 2000 (these results are also not reported). The influence of the year 2000 for immigrants reflects either a genuine change in the effect (perhaps caused by an increase in the quality of skilled immigrants through the expansion of the H1–B cap), reduced measurement error owing to larger numbers of skilled immigrants in the census, or the presence of a confounding factor correlated with increases in skilled immigrants in the 1990s. The results are not sensitive to the dropping of the 1980–1990 changes (these results are not reported).

The specifications of rows 1–6 of Table 8 are all reasonable specifications. It would seem that a one percentage point increase in the share of immigrant college graduates increases patenting by at least 8%, and perhaps by as much as 23 log points, or 21%. A one percentage point increase in the share of immigrant scientists and engineers increases patenting by at least 36 log points, or 31%, and perhaps by as much as 60 log points, or 47%. Unreported regressions suggest a one percentage point increase in the share of immigrant post–college graduates increases patenting by at least 11 log points and perhaps by as much as 47 log points (or 38%).

4 Conclusions

In this paper I have combined individual and aggregate data to demonstrate the important boost to innovation provided by skilled immigration to the United States in 1950–2000. A calculation of the effect of immigration in the 1990–2000 period puts the magnitudes of the effects in context. The 1.3 percentage point increase in the share of the population composed of immigrant college graduates increased patenting per capita by 10–26%.[14]. The 0.7 percentage point increase in the share of post–college immigrants increased patent-

[14]Between $8 \times 1.3 = 10.4$ log points= 10% and $23 \times 1.3 = 29.9$ log points= 26%.

ing by 7–28%.[15] If the lower ends of these ranges are correct, skilled immigrants have small positive spill–overs on others, while if the upper ends of the ranges are correct, the spill–overs are large. For immigrant scientists and engineers, the estimates including spill–overs imply an impact of a 0.45 percentage point increase of 15–24%,[16] but these estimates seem too large compared to the estimate without spill–overs of 9%,[17] which is a more conservative lower bound. While I find evidence that immigrants crowd out natives from certain occupations or states in the short run, I do not find evidence of crowd–out in the long run.

I find that a college graduate immigrant contributes at least twice as much to patenting as his or her native counterpart. The difference is fully explained by the greater share of immigrants with science and engineering education, implying immigrants are not innately more able than natives. Indeed, immigrants currently working as scientists and engineers are less likely to have patented recently than observably similar native scientists and engineers. Despite this, the fact that immigrants increase patenting per capita without reducing native patenting shows that their presence in the United States is beneficial to natives, assuming the immigrants would have been less innovative or less able to commercialize their innovation elsewhere or that U.S. natives benefit more from innovation and commercialization in the United States than abroad. If natives are making optimal career decisions, subsidies to induce them to enter science and engineering in greater numbers would not be beneficial even if the marginal native had higher patenting ability than immigrants in science and engineering. Policies to encourage natives to enter science and engineering are warranted only if they address obstacles to optimal decision–making, such as a lack of information about available careers, inadequate primary and secondary education or excessively high discount rates.

The results do not make clear precisely which immigration policies are appropriate to take advantage of the contributions of immigrants demonstrated in the paper. While

[15] Between $11 \times 0.7 = 7.7$ log points= 7% and $47 \times 0.7 = 32.9$ log points= 28%.

[16] Between $36 \times 0.45 = 16.2$ log points= 15% and $60 \times 0.45 = 27$ log points= 24%.

[17] $21 \times 0.45 = 9\%$.

allocating more visas based on whether the applicant has studied science or engineering may seem appealing, such a policy ignores potential benefits of immigrants without a science or engineering background. Science and engineering could also be boosted by a policy oriented towards permitting more immigrant students to remain in the United States, since such students self–select into science and engineering. Such a policy would leave students the flexibility to adjust their career plans in response to labor market developments. More research is required to investigate whether immigrants who studied in the United States contribute more to the United States than their counterparts arriving after finishing their studies.

References

[1] Aghion, Philippe and Peter Howitt. 1992. “A Model of Growth through Creative Destruction”. *Econometrica* 60 pp. 323–351.

[2] Agrawal, Ajay, Devesh Kapur and John McHale. 2007. “Brain Drain or Brain Bank? The Impact of Skilled Emigration on Poor–Country Innovation”. University of Toronto working paper.

[3] Anderson, Stuart and Michaela Platzer. n.d. “American Made: The Impact of Immigrant Entrepreneurs and Professionals on U.S. Competitiveness”. National Venture Capital Association.

[4] Baker, Michael, Dwayne Benjamin and Shuchita Stanger. 1999. “The Highs and Lows of the Minimum Wage Effect: A Time–Series Cross–Section Study of the Canadian Law”. *Journal of Labor Economics* 17 (2) pp.318–350.

[5] Barro, Robert J. and Xavier X. Sala-i-Martin. 1991. “Convergence Across States and Regions”. *Brookings Papers on Economic Activity* 1 pp. 107–158.

[6] Borjas, George J. 2006. “Do Foreign Students Crowd Out Native Students from Graduate Programs?”. In Ronald G. Ehrenberg and Paula E. Stephan eds. *Science and the University*, Madison: University of Wisconsin Press.

[7] Bottazzi, Laura and Giovanni Peri. 2003. “Innovation and spillovers in regions: Evidence from European patent data”. *European Economic Review* 47 pp. 687–710.

[8] Card, David. 2007. “How Immigration Affects U.S. Cities”. CReAM Discussion Paper No. 11/07.

[9] Card, David. 2005. “Is the New Immigration Really So Bad?” *Economic Journal* 115 (507) pp. F300–323.

[10] Card, David and John DiNardo. 2000. “Do Immigrant Inflows Lead to Native Outflows?” *American Economic Review Papers and Proceedings* 90 (2) pp. 360–367.

[11] Chellaraj, G. Keith E. Maskus and A. Mattoo. 2004. “The Contribution of Skilled Immigration and International Graduate Students to U.S. Innovation”. University of Colorado Working Paper No. 04-10.

[12] Griliches, Zvi. 1990. “Patent Statistics as Economic Indicators: A Survey”. *Journal of Economic Literature* 28 (4) pp. 1661–1707.

[13] Grossman, Gene and Elhanan Helpman. 1991a. *Innovation and Growth in the Global Economy* Cambridge: MIT Press.

[14] Grossman, Gene and Elhanan Helpman. 1991b. “Quality Ladders in the Theory of Growth”. *Review of Economic Studies* 58 pp. 43–61.

[15] Hall, Bronwyn H. 2005. “Exploring the Patent Explosion”. *Journal of Technology Transfer* 30 (1/2) pp. 35–48.

[16] Hall, Bronwyn H., Adam Jaffe and Manuel Trajtenberg. 2001. “The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools”. NBER Working Paper 8498.

[17] Hicks, Diana, Tony Breitzman, Dominic Livastro and Kimberly Hamilton. 2001. “The changing composition of innovative activity in the US – a portrait based on patent analysis”. *Research Policy* 30 pp. 681–703.

[18] Jaffe, Adam B., Manuel Trajtenberg and Rebecca Henderson. 1993. “Geographic Localization of Knowledge Spillovers as Evidence by Patent Citations”. *Quarterly Journal of Economics* 108 (3) pp. 577–598.

[19] Jones, Charles I. 1995. “Time Series Tests of Endogenous Growth Models”. *Quarterly Journal of Economics* 110 (2) pp.495–526.

[20] Kerr, William R. 2007. “The Ethnic Composition of US Inventors”. Harvard Business School Working Paper No. 08-006.

[21] Khan, B. Zorina and Kenneth L. Sokoloff. 1993. “Schemes of Practical Utility: Entrepreneurship and Innovation Among Great Inventors in the United States, 1790-1865”. *Journal of Economic History* 53 (2) pp.289–307.

[22] Marx, Matt, Deborah Strumsky and Lee Fleming. 2007. “Noncompetes and Inventor Mobility: Specialists, Stars, and the Michigan Experiment”. Harvard Business School Working Paper 07-042.

[23] Morgan, Robert P., Carlos Kruytbosch and Nirmala Kannankutty. 2001. “Patenting and Invention Activity of U.S. Scientists and Engineers in the Academic Sector: Comparisons with Industry”. *Journal of Technology Transfer* 26 pp. 173–183.

[24] Niebuhr, Annekatrin. 2006. “Migration and Innovation: Does Cultural Diversity Matter for Regional R&D Activity?” IAB Discussion Paper No. 14/2006.

[25] Peri, Giovanni. 2007. “Higher Education, Innovation and Growth”. In Giorgio Brunello, Pietro Garibaldi and Etienne Wasmer eds. *Education and Training in Europe*, Oxford: Oxford University Press.

[26] Romer, Paul M. “Endogenous Technological Change”. *Journal of Political Economy* 98 pp. S71–103.

[27] Stephan Paula E. and Sharon G. Levin. 2001. “Exceptional contributions to US science by the foreign–born and foreign–educated”. *Population Research and Policy Review* 20 pp. 59–79.

[28] Stuart, Toby E. and Olav Sorenson. "Liquidity Events and the Geographic Distribution of Entrepreneurial Activity". *Administrative Science Quarterly* 48 (2) pp. 175–201.

[29] Stuen, Eric T., Ahmed Mushfiq Mobarak and Keith E. Maskus. 2007. "Foreign PhD Students and Knowledge Creation at U.S. Universities: Evidence from Enrollment Fluctuations". University of Colorado working paper.

[30] U.S. Department of Commerce, Patent and Trademark Office. 1977. *Technology Assessment and Forecast, Seventh Report*, Washington, D.C.

[31] Zucker, Lynne G. and Michael R. Darby. 2006. "Movement of Star Scientists and Engineers and High–Tech Firm Entry". NBER Working Paper 12172.

[32] Zucker, Lynne G., Michael R. Darby, Jonathan Furner, Robert C. Liu and Hongyan Ma. 2006. "Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production". NBER Working Paper 12669.

Data Appendix

A.1 National Survey of College Graduates

The data were collected between October 2003 and August 2004 by the U.S. Bureau of the Census, on behalf of the National Science Foundation. The data consist of a stratified random sample of people reporting having a bachelor's degree or higher on the long form of the (April) 2000 census, who were under age 76 and living in the United States or its territories including Puerto Rico in the reference week of October 1, 2003. Immigrants are those born outside the United States and its territories. The United States and its territories cannot be distinguished. Missing information is imputed with a hot deck procedure, and imputed values are not flagged. More information on the data is provided at www.nsf.gov/statistics/showsrvy.cfm?srvy_CatID=3&srvy_Seri=7#fn1. The data are available at www.nsf.gov/statistics/sestat/.

A.2 Patents

I combine two patent series from the U.S. Patent and Trademark Office. The first series was compiled for me by the USPTO based on their electronic records which begin in 1963. This series is utility patents by state and year of application. Year of application is preferred to year of grant as it is a more accurate match to the time of invention. The second series (U.S. Department of Commerce 1977) is from paper–based USPTO records of patents by state and grant year 1883–1976 (application year is not available pre–1963). Grants lag applications by a median of three years between 1950 and 1963 (according to my US–wide calculations based on Lexis–Nexis), so I lead this series three years. Patents grants are also more volatile than patent applications (Hall 2005), so I smooth the series with a three year moving average. Finally, because for 1930–1960 plants and designs cannot be separated from utility patents, I leave them in for the whole series, calculate by state the average percent gap in the overlap years of the two series (18% on average), and reduce the old series by this percent. I then merge the series, using the adjusted paper series values only for pre–1963. The USPTO attributes a patent to a state according to the home address of the first–listed inventor.

I also use the NBER Patent Citation Data File, as updated by Bronwyn Hall on her website at elsa.berkeley.edu/~bhhall/bhdata.html. I weight patents by citations, and aggregate total patents as well as patents by field to the state–year level, according to filing year.

A.3 Immigration, education, age, occupation, labor force status

I use extracts from the Integrated Public Use Microdata Series for the United States Census, available at usa.ipums.org/usa/, and aggregate to the state level using the weights provided. Variables computed as shares (other than the excluded instruments) are computed as shares of the population or workers aged 18–65, and average population age is the average age of people aged 18–65. Immigrants are people born outside the United

States. I use the census–provided edurec variable to identify college graduates (16 years of education or more in the 1950–1980 censuses, and a college or higher degree in the 1990 and 2000 censuses) and high–school dropouts (11 or fewer years of education). People with post–college education are people with 17 or more years of education in the 1950–1980 censuses, and a post–college degree in 1990 and 2000. This is the highest level of education that can be distinguished for the whole 1950–2000 period. I use the 1940 census to compute lagged instruments. Alaska and Hawaii are not in the 1940 and 1950 IPUMS. The SIC codes I count as electrical engineering are 321, 322, 342, 350, 371, 372.

A.4 Other data

I use Bureau of Economic Analysis data for total state population (used to weight the regressions) and for state personal income per capita (available from 1929 onwards, unlike gross state product which is not available for my whole period). The data are available at www.bea.gov/regional/spi/.

Department of Defense procurement contracts by state are available on paper for the early years in *Prime Contract Awards by State, Fiscal Years 1951–1978*, published by the Department of Defense, OASD (Comptroller), Directorate for Information Operations and Control. The later years are available online at www.fpds.gov. Some measurement error in the attribution to states is involved, as recipient firms may subcontract the work to firms in other states. Also, in the electronic records for 1978–1983, 1986 and 1989 (of which only 1980 is relevant for the paper), the California numbers seem to be too small by a factor of 1000, so I have multiplied them by 1000. (I have obtained scanned versions of the paper documents for these years: the values for the non–problematic states and years are only approximately the same as those online, but the problematic California years are indeed about 1000 times higher than the online version.)

I obtain the land area of each state from the US. Census Bureau at www.census.gov/population/censusdata/90den_stco.txt.

Table 1: Patenting by immigrant status

	(1)	(2)	(3)	(4)	(5)	(6)
	College graduates		Post-college		Scientists/ engineers	
	Immigrant	Native	Immigrant	Native	Immigrant	Native
Any patent granted	0.019	0.009	0.036	0.013	0.063	0.049
Number patents granted	0.058	0.028	0.113	0.036	0.179	0.131
Any patent commercialized	0.012	0.006	0.021	0.008	0.037	0.030
Number patents commercialized	0.029	0.017	0.054	0.019	0.085	0.074
Share immigrant	0.134		0.157		0.236	
Observations	19,955	72,597	11,587	30,915	6644	15,715

Notes: Shares weighted with survey weights. Patents questions only asked of respondents who had ever worked. Whether a patent has been granted refers to period from October 1998 to the survey in 2003, and whether a patent has been commercialized or licensed refers to those patents granted in the same period.

Source: 2003 National Survey of College Graduates.

Table 2: Means of aggregate patents and aggregate variables affecting patenting

	1950-2000	1950	2000
Patents	2734	1343	5360
	(3931)	(1251)	(7046)
Patents/population, x100	0.023	0.018	0.035
	(0.015)	(0.011)	(0.020)
Share of population 18-65 that is:			
Immigrant, college education and above	0.016	0.004	0.035
Native, college education and above	0.136	0.062	0.200
Immigrant, post-college education	0.008	0.002	0.016
Native, post-college education	0.054	0.023	0.077
Share of workers 18-65 that are:			
Immigrant, scientists and engineers	0.004	0.001	0.009
Native, scientists and engineers	0.024	0.012	0.035
Population (millions)	9.7	6.2	12.5
	(7.8)	(4.3)	(10.1)
Age (18-65)	38.8	38.7	39.5
	(1.0)	(0.9)	(0.6)
DoD prime military procurement contracts	3221	1500	5499
(millions of nominal $)	(4379)	(1679)	(5799)
State personal income per capita (nominal $)	13,160	1504	29,845
	(11005)	(317)	(4080)
Land area (millions of square kilometers)	0.193	0.174	0.209
	(0.171)	(0.152)	(0.183)
Observations	304	49	51

Notes: Means of state-level variables, weighted by state population the year after the census. Patents and population are led by one year. Census information is not available for Alaska and Hawaii in 1950. Patents are classified by year filed.

Sources:
Education, age, occupation, nativity: U.S. Census Bureau, IPUMS decennial census microdata usa.ipums.org/usa/
Patents: U.S. Patent and Trademark Office, electronic and paper data.
State income, population: Bureau of Economic Analysis www.bea.gov/regional/spi/
Land Area: U.S. Census Bureau www.census.gov/population/censusdata/90den_stco.txt

Table 3: Patenting by field of study and field of study by immigrant status, college graduates

Field of highest degree	(1) Any patent granted	(2) Any patent commercialized	(3) Share immigrants	(4) Share natives
Computer science, math	0.017	0.012	0.080	0.036
Biological, agricultural and environment sciences	0.023	0.011	0.056	0.040
Physical sciences	0.066	0.038	0.036	0.017
Social and related sciences	0.004	0.002	0.093	0.108
Engineering	0.060	0.042	0.136	0.053
Other S&E (mainly health)	0.007	0.004	0.165	0.122
Non-S&E	0.004	0.002	0.433	0.624
All fields	0.011	0.007	1.00	1.00

Notes: Shares weighted by survey weights. "S&E" means science and engineering. Full sample (i.e. college graduates), 92,552 observations. Whether a patent has been granted refers to period from October 1998 to the survey in 2003, and whether a patent has been commercialized or licensed refers to those patents granted in the same period.

Source: 2003 National Survey of College Graduates.

Table 4: Effect of immigrant status on patenting

	(1)	(2)	(3)	(4)	(5)	(6)
		Any patent granted?			Any patent commercialized?	
College graduates	**0.0103**	**0.0010**	-0.0007	-0.0004	**0.0064**	-0.0004
	(0.0010)	**(0.0004)**	(0.0004)	(0.0003)	**(0.0008)**	(0.0003)
Pseudo-R^2	0.01	0.15	0.19	0.21	0.01	0.18
Post-college	**0.0231**	**0.0015**	0.0005	0.0006	**0.0138**	0.0002
	(0.0018)	**(0.0008)**	(0.0006)	(0.0006)	**(0.0014)**	(0.0004)
Pseudo-R^2	0.02	0.21	0.24	0.26	0.02	0.21
Scientists and engineers	**0.0140**	0.0041	**-0.0087**	**-0.0067**	**0.0069**	**-0.0046**
	(0.0039)	(0.0031)	**(0.0027)**	**(0.0026)**	**(0.0030)**	**(0.0021)**
Pseudo-R^2	0.00	0.06	0.12	0.13	0.00	0.09
Major field of highest degree	--	Y	Y	Y	--	Y
Highest degree	--	--	Y	Y	--	Y
Age, age2, sex, employed	--	--	--	Y	--	--

Notes: Marginal effect on dummy for foreign-born from weighted probits. There are 92,552 observations in the college graduate sample, 42,502 in the post-college sample and 22,359 in the scientist and engineer sample. All scientists and engineers are employed in the reference week. Highest degree categories are bachelor's, master's (including MBA), PhD and professional. There are 30 major field of study dummies (I combine the two S&E teacher training categories into one).

Table 5: Crowd-out - effect of change in immigrant skilled share on change in total skilled share

	(1)	(2)	(3)	(4)	(5)	(6)
	Weighted least squares			Instrumental variables		
Difference:	10 year	30 year	50 year	10 year	30 year	50 year
Panel A: Immigrant college+ as share of population						
Δ % Immigrant	0.51	**0.75**	**0.95**	0.26	0.56	0.79
	(0.32)	**(0.38)**	**(0.35)**	(0.30)	(0.38)	(0.47)
	[0.13]	[0.52]	[0.88]	[0.02]	[0.25]	[0.65]
R-squared	0.69	0.52	0.33	0.69	0.51	0.32
F-statistic excluded instruments	--	--	--	26	33	19
Panel B: Immigrant post-college as share of population						
Δ % Immigrant	**1.42**	**1.50**	**1.88**	0.63	1.31	**1.94**
	(0.25)	**(0.48)**	**(0.33)**	(0.59)	(0.67)	**(0.54)**
	[0.11]	[0.30]	[0.01]	[0.53]	[0.65]	[0.09]
R-squared	0.80	0.38	0.58	--	--	--
F-statistic excluded instruments	--	--	--	45	21	16
Panel C: Immigrant scientists and engineers as share of workers						
Δ % Immigrant	**1.01**	**0.79**	**1.37**	0.33	0.08	**0.76**
	(0.29)	**(0.35)**	**(0.34)**	(0.30)	(0.25)	**(0.29)**
	[0.98]	[0.56]	[0.27]	[0.03]	[0.00]	[0.42]
R-squared	0.74	0.42	0.45	--	--	--
F-statistic excluded instruments	--	--	--	5	7	9
Observations	253	151	49	253	151	49

Notes: The dependent variable is the change in the share of skilled people across periods ranging from ten to 50 years: in panel A skilled people are college graduates (as a share of the population), in panel B post-college educated (as a share of the population), in panel C scientists and engineers (as a share of workers). Regressions are weighted with weights $1/(1/pop_t+1/pop_{t-k})$, where k is equal to 10 in columns 1 and 4, 30 in columns 2 and 5, and 50 in columns 3 and 6. The instruments are the share of high school dropouts in the population at time t-k from Europe, China and India. All regressions also include change in average age and (except columns 3 and 6) year dummies. Standard errors clustered by state in parentheses. P-value of the test that the coefficient is equal to one in square brackets.

Table 6: Effect of share of immigrant college graduates on patent growth per capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Δ Log patents per capita				Δ Patents per capita		
	Weighted least squares			IV	Weighted LS		IV
Difference:	10 year	30 year	50 year	10 year	10 year	50 year	10 year
Δ % Immigrant college+	**11.4**	**11.7**	**15.6**	**17.7**	**0.389**	**0.387**	0.688
as share of population	**(4.1)**	**(3.0)**	**(4.8)**	**(7.6)**	**(0.173)**	**(0.186)**	(0.389)
				[16]			[16]
Δ % Native college+	2.1	**5.0**	**6.7**	3.3	-0.007	0.173	0.050
as share of population	(2.4)	**(2.0)**	**(2.6)**	(2.0)	(0.115)	(0.106)	(0.087)
Δ Age (average)	**0.119**	**0.147**	0.088	**0.120**	0.0023	-0.0019	0.0022
	(0.031)	**(0.049)**	(0.109)	**(0.032)**	(0.0013)	(0.0032)	(0.0014)
Δ DoD procurement	-0.031	**-0.090**	-0.063	**-0.039**	-0.0009	-0.0001	-0.0013
(log)	(0.016)	**(0.034)**	(0.074)	**(0.019)**	(0.0008)	(0.0024)	(0.0009)
Land area (log)	**0.071**	**0.207**	**0.404**	**0.078**	**0.0020**	**0.0101**	**0.0023**
	(0.012)	**(0.033)**	**(0.086)**	**(0.013)**	**(0.0005)**	**(0.0024)**	**(0.0006)**
Population 1950 (log)	**-0.049**	**-0.174**	**-0.300**	**-0.059**	**-0.0015**	-0.0076	**-0.0020**
	(0.015)	**(0.035)**	**(0.087)**	**(0.015)**	**0.0007)**	(0.0041)	**(0.0007)**
State personal income	**-0.184**	**-0.814**	**-1.483**	**-0.251**	0.0031	-0.0031	0.0001
per capita 1950 (log)	**(0.076)**	**(0.174)**	**(0.387)**	**(0.083)**	(0.0025)	(0.0114)	(0.0033)
R-squared	0.64	0.57	0.57	--	0.47	0.34	--
Observations	253	151	49	253	253	49	253

Notes: The dependent variable is the difference in (log) patents per capita across periods ranging from ten to 50 years, with a lead of one year compared to the independent variables. Weighted least squares with weights $1/(1/pop_{t+1}+1/pop_{t-k+1})$, where k the length of the difference. Regressions in columns 1,2,4, 5 and 7 include year dummies. The instrumented variable is the change in the share of immigrant college graduates; the instruments are three variables for the share of high school dropouts in the population at time t-10 from Europe, China and India. F-statistic for test of joint significance of excluded instruments in the first stage in brackets. Standard errors clustered by state in parentheses. Coefficients in columns 5-7 are multiplied by 100.

Table 7: Effect of immigrant post-college and engineer shares on patent growth per capita

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Δ Log patents per capita				Δ Patents per capita		
	Weighted least squares			IV	Weighted LS		IV
Difference:	10 year	30 year	50 year	10 year	10 year	50 year	10 year
Panel A: Immigrant post-college as share of population							
Δ % Immigrant	17.7	**21.4**	**27.4**	38.1	0.756	0.657	1.733
	(11.1)	**(8.2)**	**(11.5)**	(21.9)	(0.526)	(0.481)	(1.160)
				[16]			[16]
Δ % Native	-1.2	1.3	9.9	-2.2	-0.079	0.197	-0.125
	(3.3)	(4.4)	(6.8)	(3.6)	(0.148)	(0.289)	(0.172)
R-squared	0.63	0.52	0.52	--	0.46	0.29	--
Panel B: Immigrant scientists and engineers as share of workers							
Δ % Immigrant	**48.7**	**48.6**	**59.2**	**53.6**	**2.393**	**1.934**	**2.263**
	(20.7)	**(16.1)**	**(15.8)**	**(25.1)**	**(1.017)**	**(0.717)**	**(1.188)**
				[6]			[6]
Δ % Native	**11.8**	**20.8**	**29.5**	**11.7**	0.231	**0.866**	0.233
	(5.3)	**(6.9)**	**(7.8)**	**(5.3)**	(0.237)	**(0.287)**	(0.241)
R-squared	0.68	0.59	0.67	--	0.55	0.48	--
Observations	253	151	49	253	253	49	253

Notes: The dependent variable is the difference in (log) patents per capita across periods ranging from ten to 50 years, with a lead of one year compared to the independent variables. Weighted least squares with weights $1/(1/pop_{t+1}+1/pop_{t-k+1})$, where k is equal to the difference length. All regressions include the covariates of Table 6. The instrumented variable is the change in the share of skilled immigrants; the instruments are three variables for the share of high school dropouts in the population at time t-10 from Europe, China and India. F-statistic for test of joint significance of excluded instruments in the first stage in brackets. Standard errors clustered by state in parentheses. Coefficients in columns 4-6 are multiplied by 100.

Table 8: Determinants of patents growth per capita, specification checks for ten-year differences

	(1)	(2)	(3)	(4)
	Δ Log patents per capita			
	College+		Scientists and engineers	
Δ % Immigrant	WLS	IV	WLS	IV
1. Base specifications	**11.4**	**17.7**	**48.7**	**53.6**
(Tables 6,7)	**(4.1)**	**(7.6)**	**(20.7)**	**(25.1)**
		[16]		[6]
2. Instrument is % population which is	--	**23.1**	--	**56.5**
European, Chinese, Indian-born high		**(8.4)**		**(23.8)**
school dropouts at t-20		[10]		[10]
3. Instrument is % population which is	--	10.4	--	**58.8**
foreign-born high school dropout at t-10		(6.2)		**(22.7)**
		[28]		[15]
4. Instrument is % population which is	--	10.6	--	60.5
foreign-born high school dropout at t-20		(7.9)		(30.7)
		[17]		[5]
5. Covariates include BEA region dummies	8.1	13.6	40.8	36.3
	(4.4)	(9.5)	(24.1)	(30.3)
		[24]		[7]
6. Covariates include % workers in	**9.8**	**14.0**	**41.5**	44.5
electrical sectors 1980*year dummies	**(4.0)**	**(5.9)**	**(18.1)**	(23.1)
		[14]		[5]
7. Sample without California	7.4	6.9	17.4	7.0
(248 obs)	(3.9)	(5.0)	(12.4)	(20.1)
		[19]		[6]
8. Sample without year 2000	6.5	**12.5**	21.2	72.2
(202 obs)	(3.4)	**(4.7)**	(16.4)	(41.7)
		[23]		[5]

Notes: Each coefficient reported is the effect of a change in skilled immigrant share from a different regression. The dependent variable is the difference in (log) patents across ten years, with a lead of one year compared to the independent variables. Weighted least squares (columns 1 and 2) or instrumental variables (columns 2 and 4) with weights $1/(1/pop_{t+1}+1/pop_{t-9})$. The instruments are three variables for the share of high school dropouts in the population at time t-10 from Europe, China and India unless otherwise specified. F-statistic for test of joint significance of excluded instruments in the first stage in brackets. All regressions also include the covariates of Table 6 including the appropriate differenced share of skilled natives. Standard errors clustered by state in parentheses. 253 observations unless otherwise noted.

Appendix Table 1: Means of individual-level variables

	College graduates		Post-college		Scientists/engineers	
	Immigrant	Native	Immigrant	Native	Immigrant	Native
Highest degree:						
Bachelor's	0.57	0.65	--	--	0.43	0.68
Master's	0.28	0.26	0.66	0.74	0.40	0.26
Doctorate	0.07	0.03	0.17	0.08	0.16	0.06
Professional	0.07	0.06	0.17	0.17	0.01	0.01
Field of highest degree						
Computer science, math	0.080	0.036	0.095	0.027	0.223	0.167
Biological, agri-cultural, environ-ment science	0.056	0.040	0.064	0.030	0.092	0.093
Physical science	0.036	0.017	0.045	0.017	0.075	0.072
Social science	0.093	0.108	0.068	0.078	0.025	0.046
Engineering	0.136	0.053	0.136	0.037	0.396	0.322
Other S&E	0.165	0.122	0.199	0.158	0.068	0.058
Non-S&E	0.433	0.624	0.393	0.653	0.120	0.242
Sex (female)	0.47	0.50	0.42	0.49	0.239	0.226
Age	43.3	44.4	43.7	46.5	40.4	42.4
	(9.9)	(10.3)	(9.9)	(10.2)	(8.9)	(9.5)
Employed	0.86	0.85	0.89	0.87	1.00	1.00
Observations	19,955	72,597	11,587	30,915	6644	15,715

Notes: Means weighted with survey weights. S&E means science and engineering. "Other S&E" includes the social sciences.

Source: National Survey of College Graduates

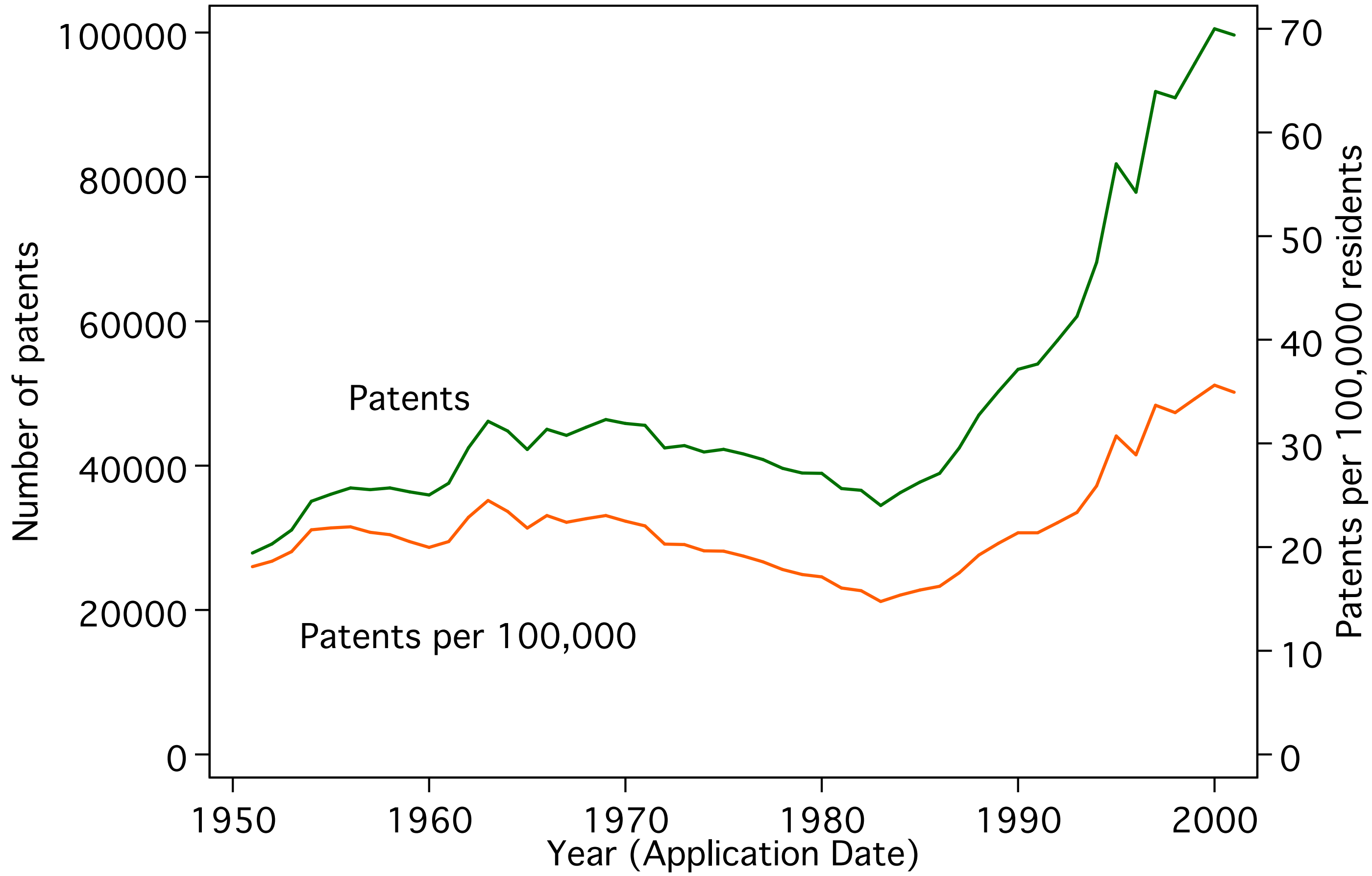
Appendix Table 2: Means of aggregate instruments for change in skilled immigrant share

	1950-2000	1950	2000
Share of population 18-65 that is:			
Immigrant, high school dropouts	0.041	0.066	0.046
Share of population 18+ that is:			
European-born, high school dropouts	0.023	0.067	0.004
Chinese-born, high school dropouts	0.0008	0.0006	0.0013
Indian-born, high school dropouts	0.0002	0.0000	0.0006
Observations	304	49	51

Notes: Means of state-level variables, weighted by state population. Census information is not available for Alaska and Hawaii in 1950.

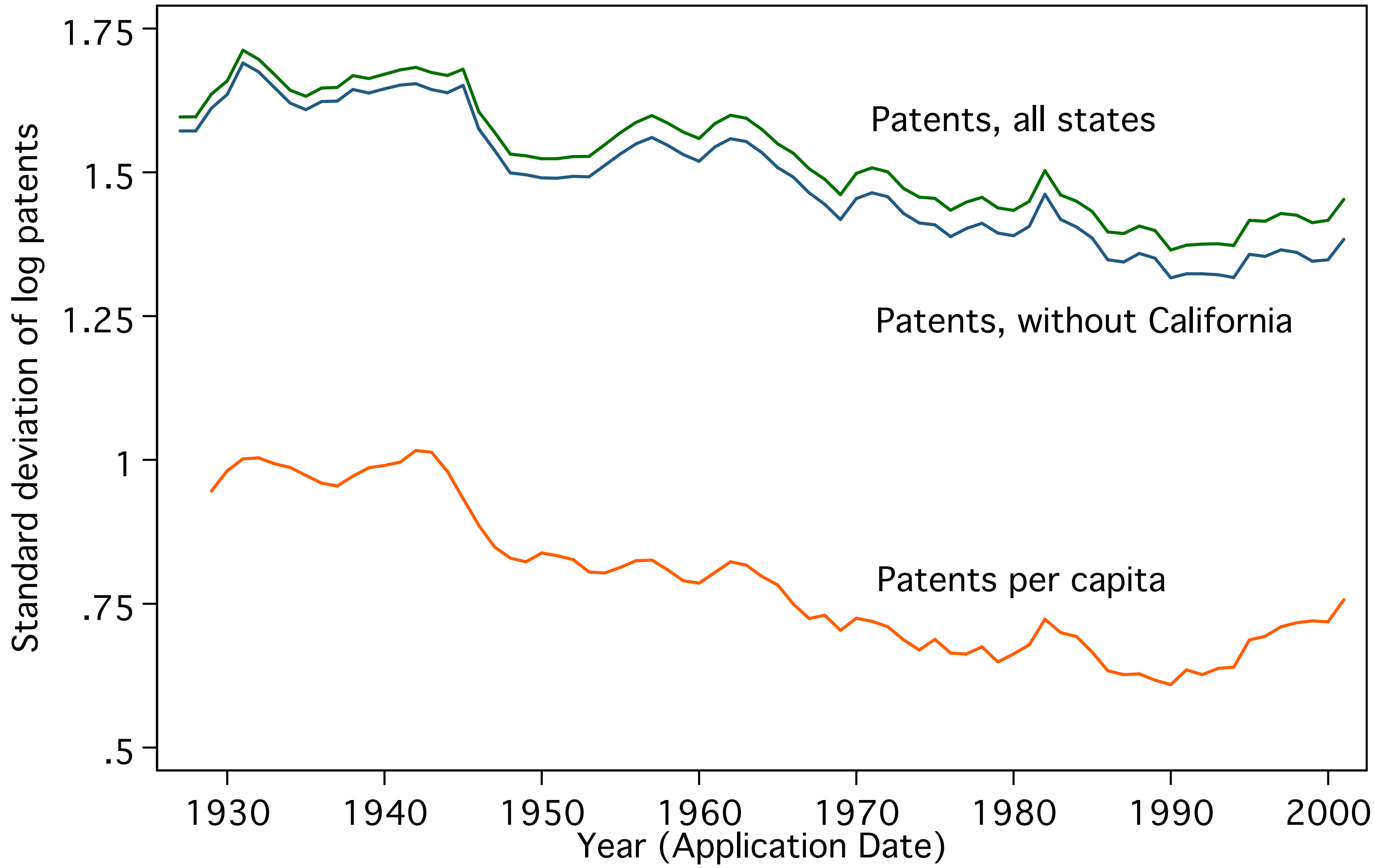
Source: U.S. Census Bureau, IPUMS decennial census microdata usa.ipums.org/usa/

Figure 1: U.S. Origin U.S. Patents 1951-2001



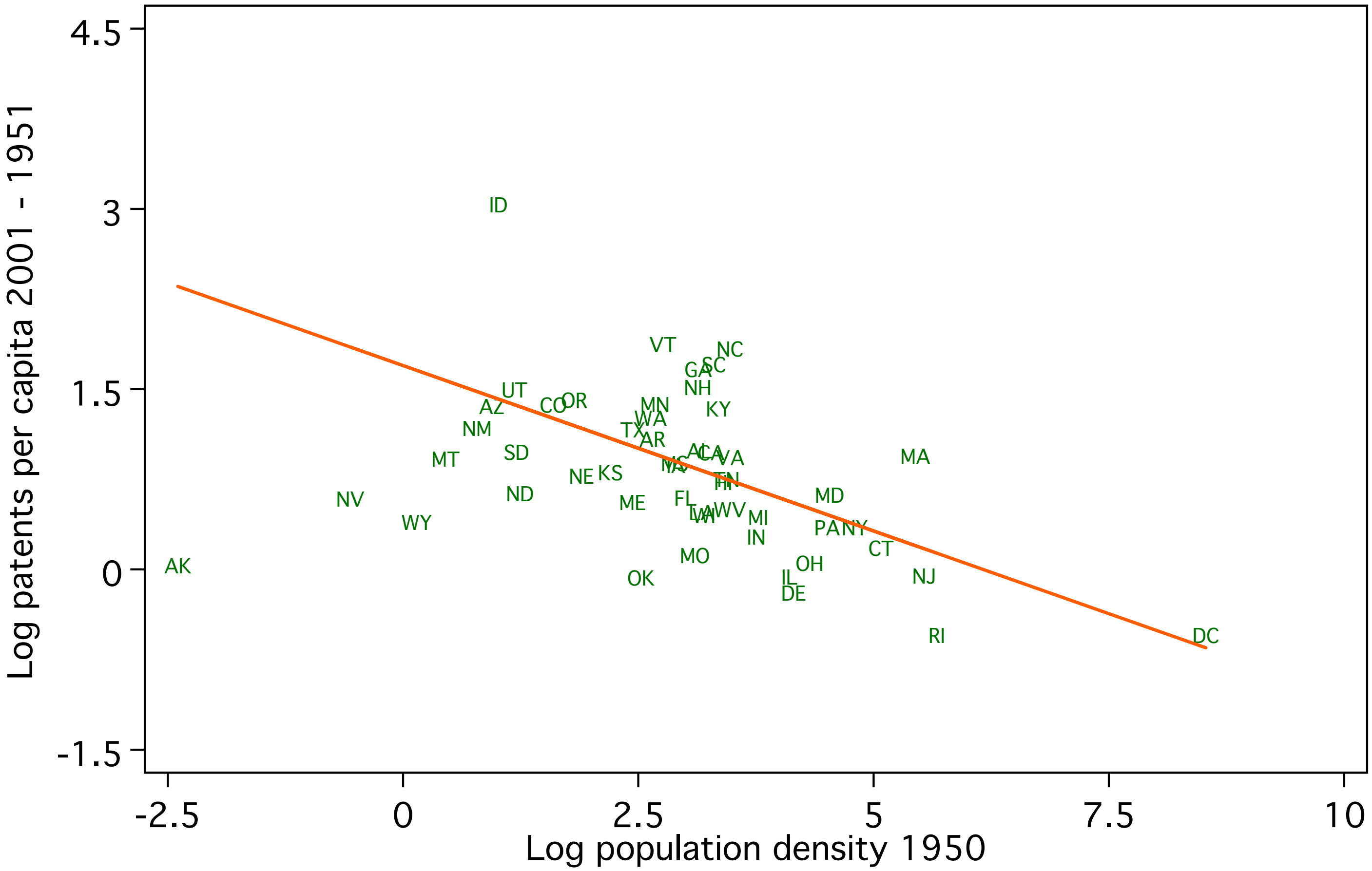
Source: USPTO, BEA and author's calculations

Figure 2: Convergence in Patenting Across States 1929-2001



Source: USPTO, BEA and author's calculations

Figure 3: Population Density as Predictor of Patent Growth



Source: USPTO and author's calculations