How Much Does Immigration Boost Innovation? *

Jennifer Hunt †

McGill University and NBER

jennifer.hunt@mcgill.ca

February 18, 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, Chad Jones, Daniel Parent, Giovanni Peri, Regina Riphahn, Eric Stuen and Dee Suttiphisal. 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 combine patent, decennial census and other data to measure the extent to which skilled immigration increased innovation in the United States from 1950–2000. I instrument 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, and consider changes of between ten and 50 years. I find that a one percentage point rise in the share of immigrant college graduates in the population increases patenting by 8–15%; the equivalent range for immigrants with post–college education is 15–33%. A one percentage point rise in the share of immigrant scientists and engineers in the workforce increases patenting by at least 41%. The effects are similar in the short and long run, and appear to be much larger than the effect of skilled natives, especially in the short run. This may be related to my finding that natives are crowded out by immigrants in the short run, but not in the long run. My analysis shows the importance of convergence among states for the evolution of patents.

Although there is a large literature studying the impact of immigration on the host country, this literature is more focused on potential costs than potential benefits. One reason for this is that the biggest potential benefits are harder to quantify than potential costs. Amongst these potential benefits are higher productivity, if there are increasing returns to scale in production; the achievement of critical mass in specialized areas of research, development and production; spill–over effects of skilled workers through externalities and production complementarities, including of the O–ring variety; increased entrepreneurship and increased innovation in science, the arts and other fields. Some tantalizing facts hint at the possible importance of these effects for the United States.[1] Twenty–six percent of U.S.–based Nobel Prize recipients from 1990–2000 were immigrants (Peri 2007), twenty–nine percent of U.S.–based U.S. patent holders had non–Anglophone names in 2000–2004 (Kerr 2007), and twenty–five percent of founders of public venture–backed U.S. companies in 1990–2005 were immigrants (Anderson and Platzer n.d.), compared to a foreign–born population of 12% in 2000. 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).

The goal of my paper is to assess the impact of skilled immigration on innovation as measured by U.S. patents. My methodology accounts for contributions through the various possible channels: innovation by the immigrants themselves, their spill–over effects and contribution to critical research and development mass, and their provision of complementary skills such as management and entrepreneurship. My analysis captures invention at companies, universities and government laboratories, and the contributions of immigrants arriving both before and after their tertiary education.

To achieve this goal, I use 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 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, and consider changes

[1] See Kremer (1993) on O–ring complementarities.

of between ten and 50 years. I also verify whether skilled immigrants crowd out skilled natives from the states (and occupations) to which they move. In so doing, I contribute to two understudied areas: the impact of immigration on innovation, and the regional determinants of innovation.

I go beyond the most closely related paper linking immigration and innovation, Peri (2007), by extending the panel, using instrumental variables, defining skilled immigration consistently across time and more broadly and by testing for crowd–out of natives. 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. The Kerr (2007) analysis of the ethnicity of inventor names relates to immigration particularly convincingly in the case of the large rise in Indian and Chinese names. In general, his analysis cannot separate the contributions of first and later generations of immigrants, however, misses the contribution of Anglophone immigrants, and is not designed to take indirect effects of immigrants into account. 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.

In a study examining the reverse phenomenon, Agrawal, Kapur and McHale (2002) find that emigration from India reduces access to knowledge in India. Another related paper is that by Niebuhr (2006), who concludes that German regions with more diverse worker nationalities (as measured by the Herfindahl) 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).

I am not aware of other academic papers examining the determinants of state–level patenting over a period of several decades (other than Peri 2007).[2] 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). The most closely related paper is Zucker and Darby (2006), who find for 1981–2004 that a Bureau of Economic Analysis (BEA) region's high–tech start–up rate is boosted by the presence of star scientists, a high

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

wage (which they view as proxying for education) and a high stock of relevant journal publications. They find these variables have no clear effect on non–university patenting, however, and they do not include region fixed effects. Zucker et al. (2006) examine the determinants of a BEA region's publications in nanotechnology, including the stock of such publications and federal funding. 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. Finally, Marx, Strumsky and Fleming (2007) and Stuart and Sorenson (2003) examine the effect of a state's enforcing non–compete laws on inventor inter–firm mobility and biotech IPOs respectively.

My work is also relevant for the macroeconomic growth literature, where the link between innovation and employment in innovation is the key to growth. AK growth models predict that the level of employment in R&D affects productivity growth (Aghion and Howitt 1992, Grossman and Helpman 1991a,b, Romer 1990), while Jones (1995) and others argue that the empirically convincing link is between the growth in such employment and productivity growth.

I find that a one percentage point rise in the share of immigrant college graduates in the population increases patenting by about 8-15%; the equivalent range for immigrants with post–college education is 15-33%. A one percentage point rise in the share of immigrant scientists and engineers in the workforce increases patenting by at least 41%. The effects are similar in the short and long run. For natives, by contrast, I am unable to find any short–run benefit of skill, and the long run beneficial effects are much smaller than those of immigrants. This could be because within each skill category natives have less education than immigrants, because immigrants are positively selected within skill category for innovative talent, or because there is bias in the effect of natives which I have not corrected with instrumental variables.

While in the short run there is some evidence that immigrants crowd out natives, either deterring natives from moving to states with skilled immigrants or deterring them from working as a scientist or engineer, in the long run there is no evidence of such crowd–out, but rather a suggestion that skilled immigrants may attract skilled natives. This is consistent with Borjas (2006), who finds that immigrants do not crowd out natives as a whole from graduate school. The absence of crowd–out means that my estimates of the benefits of immigrants are not offset by reductions in native contributions to innovation.

My quantification of immigrants' boost to innovation, in combination with Ottavanio and Peri's (2006) finding that most native education groups benefit from immigration in terms of wage, provides a more nuanced view of immigration than had hitherto prevailed. Furthermore, 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 Barro and Sala-i-Martin (1991).

1 Methodology

I use a panel of U.S. states with decennial data from 1950–2000. I have chosen the geographic dimension to supplement the time dimension as I have neither data on nor a conceptual instrument for immigrant share by firm, the usual unit of observation in patent studies. The geographic dimension is commonly used in studies of the impact of immigration on wages, where its Achilles heel is the possibility that factor price equalization makes wage responses impossible to find across states. This criticism is not relevant for a study of innovation, although I shall describe other sources of endogeneity below. I extend the period of observation back to 1950 so as to be able to distinguish long run and short run effects.[3] I do not extend it to prior decades as the years of the Great Depression and the Second World War are probably unusual.

I estimate equations in differences of lengths ranging from ten to 50 years, allowing me to judge short and long run effects purged of time–invariant state characteristics. An advantage of having a shortest difference of ten years is that results will not be as sensitive to the specification of the lags of the covariates as they would be with yearly data, and I simply define all variables contemporaneously. I estimate

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

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

where i indexes states, P is the log 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 variable for natives, and $Z_{i,1950}$ are characteristics of the state in 1950, while 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 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, as the other covariates do not appear to capture the convergence in patents per capita occurring over the time period.[4] The X covariates include 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.

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, since in some small states one company drives the time series of patenting, and I cluster standard errors by state.[6]

Although the patent and growth literatures model change in knowledge (patents) as being a function of the stock of knowledge, I elect not to include the change in the patent stock among the regressors. Its inclusion seems more appropriate for estimation pooling states without state fixed effects. I have also chosen not to use a partial adjustment

[4]See Barro and Sala-i-Martin (1991) on convergence in state personal income per capita over time.

[5]Methodologically it might be preferable to use electrical engineering employment in 1950, but it is tiny in most states until 1980.

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

model, which seems better suited to shorter run analysis.[7]

Equation (1) 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.[8] I use several sets of instruments to address these problems for foreign skilled workers. To instrument $\Delta I^S = I_t^S - I_{t-k}^S$, I use I_{t-k}^{HSD}, 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 (1). A variant on 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. I also use the wage return to college or more at $t-k$. This should not directly affect innovation growth, but high levels should attract skilled workers, particularly mobile new immigrants.[9]

Finally, I use a different proxy for cultural amenities likely to attract skilled workers, particularly from Europe and particularly those mathematically inclined (and therefore likely musically inclined): information on the presence of symphony orchestras in the state. I include dummies for the presence of any of ten types of symphony orchestra in 2006, as well as the number of each type (the categories, principally budgetary, are listed in the Appendix Table). It would be preferable to have decadal data, but these are not available. For ΔN^S (native skilled workers), I have experimented unsuccessfully with lagged college enrollments as an instrument.

A different set of concerns is related to the potential crowd–out of natives. Natives may choose not to enter careers in science and engineering, or to work less, owing to

[7] 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.

[8] There is definitely 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.

[9] I use the coefficient on having college or more from a log weekly wage regression also including sex and a quartic in age.

competition from immigrants whose comparative advantage is in less language–intensive and less institution–specific occupations. Any drop in native innovators must be taken into account when calculating the net benefit of immigrants. A more complex concern is that native innovators forgo migration to states with many immigrant innovators. If by doing so native innovators forgo an increase in productivity, this must be taken into account when calculating the net benefit of immigrants. For this to be logical, natives must be forgoing migration because of personal distaste for foreigners, since if they are forgoing a productivity increase they must also be forgoing 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

$$\Delta S = a + b\Delta I^S + f\Delta Age + \Delta\nu, \qquad (2)$$

where S is the share of the population or workforce (aged 18–65) composed of skilled natives and immigrants, and Age is the average age of the state's population between 18 and 65. 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 $\hat{b} > 1$ is also possible and would indicate that skilled natives were attracted to states with many skilled immigrants, perhaps because of positive spill–overs. Measurement error could also caused $\hat{b}$ to be less than one.

2 Data and Descriptive Statistics

The patent data used in most of the analysis come from the U.S. Patent and Trademark Office. I merge a series based on their 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 patents in the four states with the most patents in 1950: New York, California, Illinois and Ohio. The four states have similar numbers of patents in 1950, but by 2000 California has pulled away from the others. Figure 2 shows the group of four states with the most patents in 2000, a group to which New York and California also belong. By 2000 Texas and Massachusetts had overtaken Illinois and Ohio, and Texas had drawn level with New York. Figure 3 shows the states with the fastest patent growth over the period. These are states with small populations, initially small numbers of patents (note the different y–axis scale) and high population growth: Arizona, Idaho, Nevada and Utah. Idaho's growth is driven by one semi–conductor company, Micron Technology Inc., founded in 1978, which was granted 1304 patents in 2000 and was the seventh–ranked company in this regard.

In Figure 4 I use patent data from 1929 to 2000 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 1990–2000, 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 from 1975. These data permit patents to be weighted by citations to yield a quality–adjusted patent measure. The data are imperfect for my decadal analysis, however: most patents filed in 2000 have not yet been awarded, so I must use 1997 patents to match to the 2000 census; and I can begin the analysis at best with 1970, a year for which most citations are not recorded.

To compute the shares of the population in various education and occupation classes, to divide these into immigrant and foreign, 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 is the highest education level that can be measured consistently throughout 1950–2000. 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. 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. Finally, I use data from the League of American Orchestras for symphony orchestras by state by budget category for 2006.

Full details on the data construction are given in the Data Appendix, while the variable means, weighted by population, are reported in Table 1. 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 share 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 by six to 0.5%, while the native share rose from 1.2% to 3.8%. The Appendix Table contains the means of the variables used as instruments.

3 Results

Before estimating the determinants of patenting, I establish whether crowd–out of natives from states or occupations with many skilled immigrants should be a concern.

3.1 Crowd–out

To test for crowd–out, I estimate equation (2). The results with college or more as an indicator of skill are reported in Panel A of Table 2. Column 1 shows that with weighted least squares and first (i.e. ten year) differences, it appears that 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. However, as I increase the length of the differences I am using, evidence of crowd–out disappears: the coefficient is 0.75 for third (i.e. 30 year) differences in column 2, and 0.95 for fifth differences in column 3. In columns 4–6 I report the corresponding instrumental variables results, using the shares of European, Chinese and Indian dropouts as instruments (the bracketed F–statistic for their joint significance in the first stage suggests they are powerful). The coefficients are smaller – and in the case of first differences significantly different from one - but also increase as the difference length increases, to 0.79 in column 6.

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 coefficient falls then rises with the length of the difference, with the point estimate indicating (statistically insignificant) crowd–out for third differences. The instruments used in columns 4–6 are not as powerful in the first stage as for the other skill measures, except for fifth differences, where the crowd–out coefficient is 0.83.

In summary, it appears that in the short–run there is some crowd–out of college–educated natives by college–educated immigrants, but that particularly in the longer run and for the other skill groups the opposite occurs: skilled natives are attracted to skilled immigrants. I cannot correct for crowd–out when analyzing the determinants of innovation, but the results mean that the short–run benefits of college immigrants may be overestimated, while other benefits may be underestimated.

3.2 Determinants of patents

The evolution of a state's patents over the period 1950–2000 is strongly related to conditions in 1950, as would be expected given the convergence depicted in Figure 4. This is vividly illustrated by Figure 5, which measures the difference in log patents per capita over the period on the y–axis. The x–axis is the population density of the state in 1950 (population divided by land area in square kilometers). The 1950 population density explains 33% 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.[10]

I use these observations to inform my choice of specifications in Table 3, where I examine the effect on patenting of the share of skilled people, without yet distinguishing between immigrants and natives. As in the figure, I focus on the fifth difference. Panel A contains the coefficients from regressions with skill represented by the share of the population that has a college degree, panel B for the share with post–college, and panel C for the share of workers who are scientists or engineers. The coefficients in column 1, where there are no other covariates, are negative for college and post–college, and they remain negative in column 2 when further covariates are added (the changes in log DoD spending and average age). The column 1 coefficient for scientists and engineers is positive but not quite significant at 11.4, and falls to an insignificant 7.5 in column 2.

It is possible that education has no general beneficial effect on patenting, and that the benefit of scientists and engineers is reflected by the insignificant coefficient of 7.5 in column 2. The negatively signed effects of education are robust to various lengths of differences and to the use of instrumental variables (these results are not reported). However, the elasticity implied by the science and engineering coefficient is 0.21 (using the mean share of scientists and engineers from Table 1, 2.8%), which seems implausibly small. If all patents were filed by scientists and engineers, one might expect the elasticity

[10]Washington D.C.'s decline is influenced by a sharp decline in NASA patenting in the 1970s in addition to its severe population decline. The location of a patent is determined by the home address of the first inventor, which for D.C. means that suburbanization reduces patenting.

to be approximately unity (implying a coefficient of 36).

Adding as controls the log of state land area and 1950 log population, which together give 1950 log population density, and the 1950 log state income per capita, in column 3, renders all three skill coefficients positive and significant. If this is the appropriate specification, it means that there exists a factor driving convergence whose omission obscures the benefits of a skilled population (even in instrumental variables estimation).

The coefficient of 31.4 in column 3 panel C indicates that a one percentage point increase in the share of the workforce that is scientists and engineers is associated with a 31.4 log point (27%) increase in patenting, or an elasticity of 0.88, which seems reasonable. The coefficient of 7.2 in panel A indicates that a one percentage point increase in the share of the population with a college degree is associated with an increase in patents of 7.2%. The fact that many college educated do not work will tend to make this coefficient smaller than the scientist and engineer coefficient, as will the fact that the share of patenters is likely to be lower amongst college graduates generally than amongst scientists and engineers. On the other hand, the indirect effects of college graduates could be important. A one percentage point increase in post–college is associated with an increase in patenting of 12.5%. I check the sensitivity of these results to the addition of seven dummies for BEA regions in column 4: while the coefficient is little affected for scientists and engineers, the coefficients are reduced but still statistically significant for the other two skill groups. I elect to control for 1950 conditions in all subsequent analysis.

In Table 4 I distinguish between natives and immigrants with a college degree, and show the results of changing the length of the difference using weighted least squares. The coefficients on the share of immigrant college graduates are positive and significant: a one percentage point increase in the share of the population composed of immigrant college graduates is associated with an 11% increase in patenting for first, second and third differences, and 13.7–15.4% increases for fourth and fifth differences.

By contrast, there is no significant effect of native college graduates at short differences. The point estimate increases as the difference length increases, however, and for fifth differences (column 5) the coefficient is a significant 5.9. 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). Crowd–out is expected to operate principally by reducing the quantity of skilled natives, rather than the coefficient on their share, but if the composition of skilled natives shifts away from science and engineering, their coefficient could also be attenuated, possibly explaining why the native effect is larger for long differences, where crowd–out is less important. The impact of natives could be smaller than that of immigrants at long differences either because more college immigrants have post–college education, or because immigrants are positively selected for innovative talent.

Older populations appear to be more innovative. This may reflect the importance of management or other skills complementary to innovation.[11] As suggested by time series work in Griliches (1990), Department of Defense spending lowers patenting, presumably in part because military innovation 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 5 with post–college education (panel A) and science and engineering occupation (panel B) as measures of skill. The coefficients for immigrant post–college are 21–22 for 10, 20 and 30 year differences (columns 1–3), and 25–29 for longer differences (columns 4 and 5). These estimates are higher and less precisely estimated than for college graduates in Table 4. As post–college immigrants represent about half the college and above immigrants (see Table 1), if the coefficient for post–college were double the coefficient for college and above it would suggest that only the post–college are contributing to innovation. This is approximately the case. The coefficients for native post–college educated are never statistically significant though the point estimates are higher for the longer differences.

In panel C the coefficients are significant at all difference lengths for both immigrants

[11]Innovation rises monotonically with age in unreported regressions with three variables representing age shares.

and natives. The coefficients are similar across columns 1–4 for immigrants (64–69, meaning a response of 49–52% to a one percentage point increase in the share of immigrant scientists and engineers), and slightly higher at fifth differences. The coefficients increase with difference length for natives (from 12.1 to 26.3), nevertheless indicating that natives have less impact on invention than immigrants even in the long run.

The next step is to see whether the conclusions for immigrants hold up in instrumental variables regressions. The crowd–out results of Table 2 suggested that to avoid overstating benefits of immigrants, results from longer differences should be preferred. Although the immigrant least squares results of Tables 4 and 5 did not differ much by difference length, it would be preferable to focus on longer differences for the instrumental variables analysis in case the true causal effect does differ. With the exception of the orchestra instruments, however, my instruments are good predictors of the change in skilled immigrant shares only for first and second differences (for which results are similar). I therefore present the results of instrumental variables for first differences in Table 6, and hope that since the short and long–run correlations are similar, the short and long–run causal effects are similar (which is the case when the orchestra instruments are used). I report only the coefficient on the change in skilled immigrant share.

In the first row of Table 6 I use as instruments the shares of the population composed of European, Chinese and Indian high school dropouts. The coefficients for the three skill groups are all higher than their least squares equivalents (though not statistically significantly so). This is surprising as the coefficients were expected to be biased up. There are two possible explanations. One is that measurement error is an important factor, a possibility mooted by Card and DiNardo (2000) in a similar context. The other is 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. The F–statistic reported in square brackets indicates that the instruments are powerful in the first stage for college and post–college, but not for engineers and scientists.

In the second row I use the share of all immigrant high school dropouts, and its square, as instruments. These instruments are powerful in the first stage for all three skill groups. With this instrument set, the point estimates are very similar to the weighted least squares results in Tables 4 and 5 (but the standard errors are larger, so only the scientists and engineers coefficient is significant). In the third row I use the wage return to college as an instrument, which is strongly significant in the first stage for college and post–college educated, but not for scientists and engineers.[12] The coefficients for college and post–college are similar to those of the first row, while the coefficient for scientists and engineers is anomalously high at 104.2. In the final row I use the set of 14 variables describing symphony orchestras as instruments, which are strongly significant in the first stage for all three skill groups. The estimated coefficients are similar to the first row for college and post–college, and similar to weighted least squares for scientists and engineers.

The results of Table 6 suggest that the coefficients range from 11–18 for the college–educated, from 26–52 (or 23–42%) for the post–college, and from 70–104 for scientists and engineers, but with the coefficient of 70 (53%) more reliable.

In Table 7 I present the results of various specification checks for first differences, both weighted least squares and instrumental variables, using the shares of European, Chinese and Indian dropouts as instruments (this instrument is not very effective for scientists and engineers, but I present the results regardless). In the first row I reproduce the baseline results from Tables 4–6. In the second row, I add seven BEA region dummies. This reduces the coefficients to 70–90% of the magnitude of the first row and increases the standard errors, rendering all the IV estimates insignificant. In the third row, 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 first row, this time 74–86% of the baseline magnitudes, but all are statistically significant.

In the fourth row I investigate the influence of California in the baseline specification by dropping that state. This reduces the estimates by half, in some cases. Finally, I assess the robustness to dropping the 1990–2000 differences (while retaining California),

[12] Oddly, the wage return to being an engineer, relative to all other workers, is less significant in the first stage than the wage return to college.

using the baseline specification. This causes the weighted least squares coefficients to become much smaller and far from significant, with values of 4.5, 3.5 and 16.6 for college, post–college, and scientists and engineers respectively. However, for the college educated, for whom the instruments remain strongly significant in the first stage, the larger IV estimate of 8.3 is statistically significant. For the two other skill groups, the instruments are not powerful and the IV estimate, while much larger than the weighted least squares, is statistically insignificant. 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). It is not clear whether the influence of the year 2000 for immigrants reflects a genuine change in the effect (perhaps caused by an increase in the quality of skilled immigrants), 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).

Considering the results of both Tables 6 and 7 (and ignoring the least squares results without 2000), 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 15%, but probably not by as much as the 18% baseline IV result. A one percentage point increase in the share of immigrant post–college graduates increases patenting by at least 15% and perhaps by as much as 39 log points (or 33%), but probably not by as much as the 52 log points of Table 6. A one percentage point increase in the share of immigrant scientists and engineers increases patenting by at least 51 log points, or 41%. This latter magnitude is somewhat higher than the 36 log point effect that would be caused by all scientists and engineers if the elasticity were unity (a 1% increase in scientists and engineers increased patents by 1%). It is hard to pick the upper bound for scientists and engineers, but the Table 6 coefficient of 104.2 (71%) is likely above the true magnitude.

I found similar results when I used citation–weighted patent counts for the 1970–2000 period, and I found that the impact of skilled immigrants was 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.[13]

4 Conclusions

In this paper I have demonstrated the important boost to innovation per capita 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 1990–2000 increase from 2.2% to 3.5% in the share of the population composed of immigrant college graduates increased patenting by at least $8 \times 1.3 = 10.4\%$, and perhaps by as much as 18%. The increase in the share of post–college immigrants from 0.9% to 1.6% increased patenting by at least 10.5% and perhaps by as much as 24%. The increase from 0.30% to 0.55% in the share of workers who are immigrant scientists and engineers increased patenting by at least 13% but probably by less than 23%. While I find evidence for the crowding–out of natives in the short run, in the long run there is evidence for the reverse: that skilled natives are attracted to states or occupations with skilled immigrants.

The results hint that skilled immigrants innovate more than their native counterparts, especially if they are scientists or engineers. If correct, the result could reflect higher education of immigrants within skill categories, or positive selection of immigrants in terms of ability to innovate. However, the effect of natives is not as well identified econometrically as the effect of immigrants.

The paper also provides insight into issues related to the macro growth literature. I have shown the importance of convergence in patenting across states in the 1950–2000 period, which is likely to be an important element in the convergence of state personal income per capita. My results also quantify the link between innovation and its labor inputs, a crucial element of growth models.

[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.

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. 2005. "Is the New Immigration Really So Bad?" *Economic Journal* 115 (507) pp. F300–323.

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

[10] 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.

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

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

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

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

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

[16] 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.

[17] 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.

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

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

[20] Kremer, Michael. 1993. "The O–Ring Theory of Economic Development". *Quarterly Journal of Economics* 108 (3) pp. 551–576.

[21] 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.

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

[23] Ottaviano, Gianmarco I.P. and Giovanni Peri. 2006. "Rethinking the Effects of Immigration and Wages". U.C. Davis working paper.

[24] 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.

[25] Popp, David, Ted Juhl and Daniel K.N. Johnson. 2003. "Time in Purgatory: Determinants of the Grant Lag for U.S. Patent Applications". NBER Working Paper 9518.

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

[31] 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 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 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.2 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 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. People with college education or more are people with 16 years of education or more in the 1950–1980 censuses, and a college or higher degree in the 1990 and 2000 censuses. 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. Alaska and Hawaii are not in the 1950 IPUMS.

A.3 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 also requested and received 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.

The membership directory for the League of American Orchestras is available at www.americanorchestras.org/utilities/orchestra_members.html (when I accessed it on February 1, 2007 the 2006 directory was posted; now the 2007 directory is posted). Although the League of American Orchestras was founded in 1942, membership data (other than the founding membership) only exist from the 1970s onwards (furthermore, in the early years of the League, not all symphony orchestras were members). Operas and their orchestras are not members. Member college orchestras are largely those of music schools.

Table 1: Means of patents and variables affecting patenting

	1950-2000	1950	2000
Patents	2752	1694	5237
	(3755)	(1625)	(6647)
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	0.003	0.001	0.005
Native, scientists and engineers	0.025	0.012	0.038
Population (millions)	9.5	6.1	12.3
	(7.7)	(4.3)	(9.9)
Age (18-65)	38.8	38.7	39.5
	(1.0)	(0.9)	(0.6)
DoD prime military procurement contracts	3215	1499	5479
(millions of nominal $)	(4372)	(1681)	(5789)
State personal income per capita (nominal $)	13,176	1504	29,843
	(11007)	(317)	(4086)
Land area	0.193	0.172	0.208
(millions of square kilometers)	(0.170)	(0.150)	(0.183)
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. 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 2: Crowd-out - effect of change in immigrant skilled share on total skilled share

	(1)	(2)	(3)	(4)	(5)	(6)
	Weighted least squares			Instrumental variables		
	1st diffs	3rd diffs	5th diffs	1st diffs	3rd diffs	5th diffs
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	--	--	--	26	33	19
instruments (p-value)				(0.000)	(0.000)	(0.000)
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	0.78	0.37	0.58
F-statistic excluded	--	--	--	45	21	16
instruments (p-value)				(0.000)	(0.000)	(0.000)
Panel C: Immigrant scientists and engineers as share of workers						
Δ % Immigrant	**0.95**	0.66	**1.50**	0.38	0.00	**0.83**
	(0.41)	(0.45)	**(0.46)**	(0.43)	(0.26)	**(0.32)**
	[0.90]	[0.45]	[0.28]	[0.15]	[0.00]	[0.58]
R-squared	0.69	0.34	0.31	0.68	0.29	0.25
F-statistic excluded	--	--	--	4	9	16
instruments (p-value)				(0.014)	(0.000)	(0.000)
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 40 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 1 in columns 1 and 4, 3 in columns 2 and 5, and 5 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 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 3: Effect of skilled share on patent growth per capita, fifth difference (2000-1950)

	(1)	(2)	(3)	(4)
Panel A				
Δ % College+	-2.6	-0.6	**7.2**	**5.9**
as share of population	(2.4)	(1.9)	**(2.3)**	**(2.2)**
R-squared	0.03	0.35	0.58	0.71
Panel B				
Δ % Post-college	-7.2	-4.5	**12.5**	**8.3**
as share of population	(4.0)	(3.2)	**(4.3)**	**(4.0)**
R-squared	0.06	0.37	0.55	0.68
Panel C				
Δ % Scientists and engineers	11.4	7.5	**31.4**	**29.7**
as share of workers	(5.7)	(7.2)	**(6.8)**	**(6.1)**
R-squared	0.04	0.36	0.66	0.79
Δ Age, Δ Log DoD spending	--	Yes	Yes	Yes
Log land area, 1950 log population, 1950 log personal income per capita	--	--	Yes	Yes
BEA region dummies (7)	--	--	--	Yes
[p-value for joint significance]				[0.00]

Notes: The dependent variable is the difference in log patents per capita between 2000 and 1950. The three panels represent three different sets of regressions, each with 49 observations. Weighted least squares with weights $1/(1/pop_{2000}+1/pop_{1950})$. Standard errors clustered by state in parentheses.

Table 4: Effect of share of immigrant college graduates on patent growth per capita

	(1) 1st diffs	(2) 2nd diffs	(3) 3rd diffs	(4) 4th diffs	(5) 5th diffs
Δ % Immigrant college+	**11.2**	**11.1**	**11.3**	**13.7**	**15.4**
as share of population	**(3.7)**	**(2.8)**	**(2.7)**	**(3.2)**	**(4.3)**
Δ % Native college+	1.1	2.3	**4.7**	**6.0**	**5.9**
as share of population	(2.2)	(1.9)	**(1.8)**	**(2.0)**	**(2.5)**
Δ Age (average)	**0.105**	**0.126**	**0.149**	0.097	0.092
	(0.029)	**(0.036)**	**(0.047)**	(0.068)	(0.109)
Δ DoD procurement	**-0.036**	**-0.070**	**-0.088**	-0.084	-0.058
(log)	**(0.018)**	**(0.025)**	**(0.035)**	(0.048)	(0.070)
Land area (log)	**0.065**	**0.119**	**0.205**	**0.287**	**0.376**
	(0.011)	**(0.021)**	**(0.030)**	**(0.041)**	**(0.077)**
Population 1950 (log)	**-0.049**	**-0.104**	**-0.164**	**-0.223**	**-0.286**
	(0.014)	**(0.022)**	**(0.033)**	**(0.053)**	**(0.079)**
State personal income per	**-0.219**	**-0.555**	**-0.913**	**-1.397**	**-1.580**
capita 1950 (log)	**(0.075)**	**(0.125)**	**(0.172)**	**(0.252)**	**(0.360)**
R-squared	0.64	0.71	0.63	0.64	0.62
Observations	253	202	151	100	49

Notes: The dependent variable is the difference in log patents per capita across periods ranging from ten to 50 years. Weighted least squares with weights $1/(1/pop_t+1/pop_{t-k})$, where k is equal to 1 in column 1, 2 in column 2, 3 in column 3, etc. Regressions in columns 1-4 include year dummies. Standard errors clustered by state in parentheses.

Table 5: Effect of immigrant post-college and engineer shares on patent growth per capita

	(1) 1st diffs	(2) 2nd diffs	(3) 3rd diffs	(4) 4th diffs	(5) 5th diffs
Panel A: Immigrant post-college as share of population					
Δ % Immigrant	**21.6**	**21.6**	**20.7**	**25.2**	**28.5**
	(10.2)	**(8.0)**	**(7.6)**	**(8.8)**	**(10.6)**
Δ % Native	-0.8	-4.5	0.5	5.7	7.7
	(2.8)	(3.2)	(0.1)	(4.8)	(6.2)
R-squared	0.64	0.69	0.59	0.59	0.57
Panel B: Immigrant scientists and engineers as share of workers					
Δ % Immigrant	**68.5**	**64.3**	**66.8**	**67.6**	**82.3**
	(15.2)	**(11.5)**	**(10.1)**	**(10.3)**	**(16.8)**
Δ % Native	**12.1**	**16.0**	**20.1**	**22.3**	**26.3**
	(4.7)	**(5.2)**	**(5.7)**	**(5.8)**	**(6.4)**
R-squared	0.68	0.74	0.67	0.68	0.72
Observations	253	202	151	100	49

Notes: The dependent variable is the difference in log patents per capita across periods ranging from ten to 50 years. Weighted least squares with weights $1/(1/pop_t+1/pop_{t-k})$, where k is equal to 1 in column 1, 2 in column 2, 3 in column 3, etc. All regressions include the covariates of Table 4. Standard errors clustered by state in parentheses.

Table 6: Determinants of patent growth per capita, first differences, instrumental variables

Coefficient on Δ % Immigrant when excluded instruments are:	(1) College+	(2) Post-college	(3) Scientists/engineers
% population which is European, Chinese, Indian-born high school dropouts at t-1	**17.9**	**43.0**	**87.6**
	(6.6)	**(19.3)**	**(23.8)**
	[16]	**[16]**	**[5]**
% population which is foreign-born high school dropout at t-1	10.8	26.5	**70.4**
	(5.9)	(16.4)	**(25.3)**
	[28]	[26]	**[23]**
Return to college or more at t-1	**17.1**	**52.4**	**104.2**
	(8.3)	**(25.9)**	**(47.8)**
	[27]	**[25]**	**[11]**
Information on numbers of orchestras in different budget categories in 2006	**16.6**	**36.1**	**69.1**
	(3.7)	**(12.7)**	**(11.1)**
	[22]	**[29]**	**[74]**

Notes: Each coefficient reported is from a different regression. The dependent variable is the difference in log patents per capita across ten years. Weighted instrumental variables with weights $1/(1/pop_t+1/pop_{t-1})$. 253 observations. All regressions also include the covariates of Table 4. Standard errors clustered by state in parentheses. F-statistic for test of joint significance of excluded instruments in the first stage in brackets. The orchestra instruments comprise a dummy for whether the state has any orchestra in each of the ten categories, as well as the number it has in each category (see Appendix Table for the categories and means).

Table 7: Determinants of patents per capita, specification checks for first differences

	(1)	(2)	(3)	(4)	(5)	(6)
	College+		Post-college		Scientists/engineers	
Δ % Immigrant	WLS	IV	WLS	IV	WLS	IV
Base specifications	**11.2**	**17.9**	**21.6**	**43.0**	**68.5**	**87.6**
(from Tables 4,5,6)	**(3.7)**	**(6.6)**	**(10.2)**	**(19.3)**	**(15.2)**	**(23.8)**
		[16]		**[16]**		**[5]**
With BEA region	**7.9**	15.4	15.4	38.5	**60.6**	72.8
dummies	**(3.9)**	(8.6)	(9.3)	(25.2)	**(22.4)**	(41.2)
		[24]		[19]		[6]
With % workers in	**9.6**	**14.3**	**16.8**	**32.0**	**57.9**	**74.5**
electrical sectors 1980 *	**(3.6)**	**(5.2)**	**(8.4)**	**(13.6)**	**(13.6)**	**(23.4)**
year dummies		**[14]**		**[13]**		**[4]**
Without California	**8.1**	9.1	11.6	15.1	**43.7**	39.2
(248 obs)	**(3.7)**	(5.1)	(7.6)	(12.2)	**(16.4)**	(41.0)
		[19]		[20]		[14]
Without year 2000	4.5	**8.3**	3.5	34.6	16.6	50.6
(202 obs)	(2.9)	**(4.0)**	(6.7)	(18.4)	(13.9)	(39.0)
		[23]		[7]		[2]

Notes: Each coefficient reported is from a different regression. The dependent variable is the difference in log patents across ten years. Weighted least squares (columns 1-3) or instrumental variables (columns 4-6) with weights $1/(1/pop_t+1/pop_{t-1})$. The instruments are the share of high school dropouts in the population at time t-1 from Europe, China and India. All regressions also include the covariates of Table 4. Standard errors clustered by state in parentheses. 253 observations unless otherwise noted.

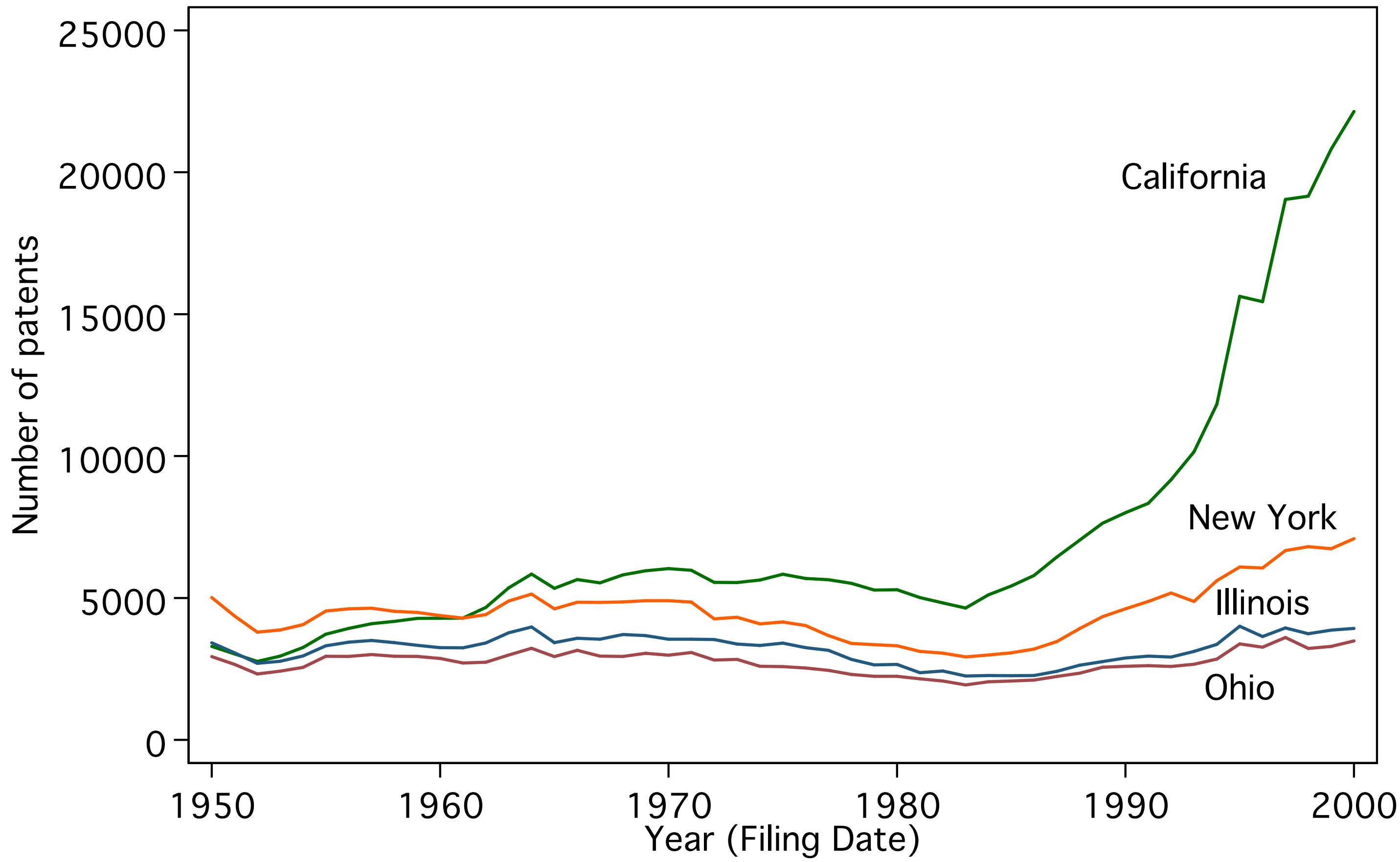
Appendix Table: Means of instruments for change in skilled immigrant share

	1950-2000	1950	2000
Share of population 18-65 that is:			
Immigrant, high school dropouts	0.041	0.067	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
Return to college or more	0.45	0.38	0.51
(vs all lower education levels)	(0.11)	(0.20)	(0.06)
Number of symphony orchestras in category 2006:			
1. Budget > $14,750,000	0.95	--	--
	(0.79)		
2. Budget $5,550,000-$14,750,000	0.90	--	--
	(1.21)		
3. Budget $2,800,000-$5,550,000	1.05	--	--
	(1.01)		
4. Budget $1,700,000-$2,800,000	1.33	--	--
	(1.83)		
5. Budget $880,000-1,700,000	2.15	--	--
	(1.77)		
6. Budget $455,000-$880,000	3.97	--	--
	(3.18)		
7. Budget $130,000-$455,000	4.47	--	--
	(2.47)		
8. Budget <$130,000	8.41	--	--
	(6.59)		
9. Youth	9.04	--	--
	(7.64)		
10. College (mostly music schools)	2.84	--	--
	(2.51)		
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.

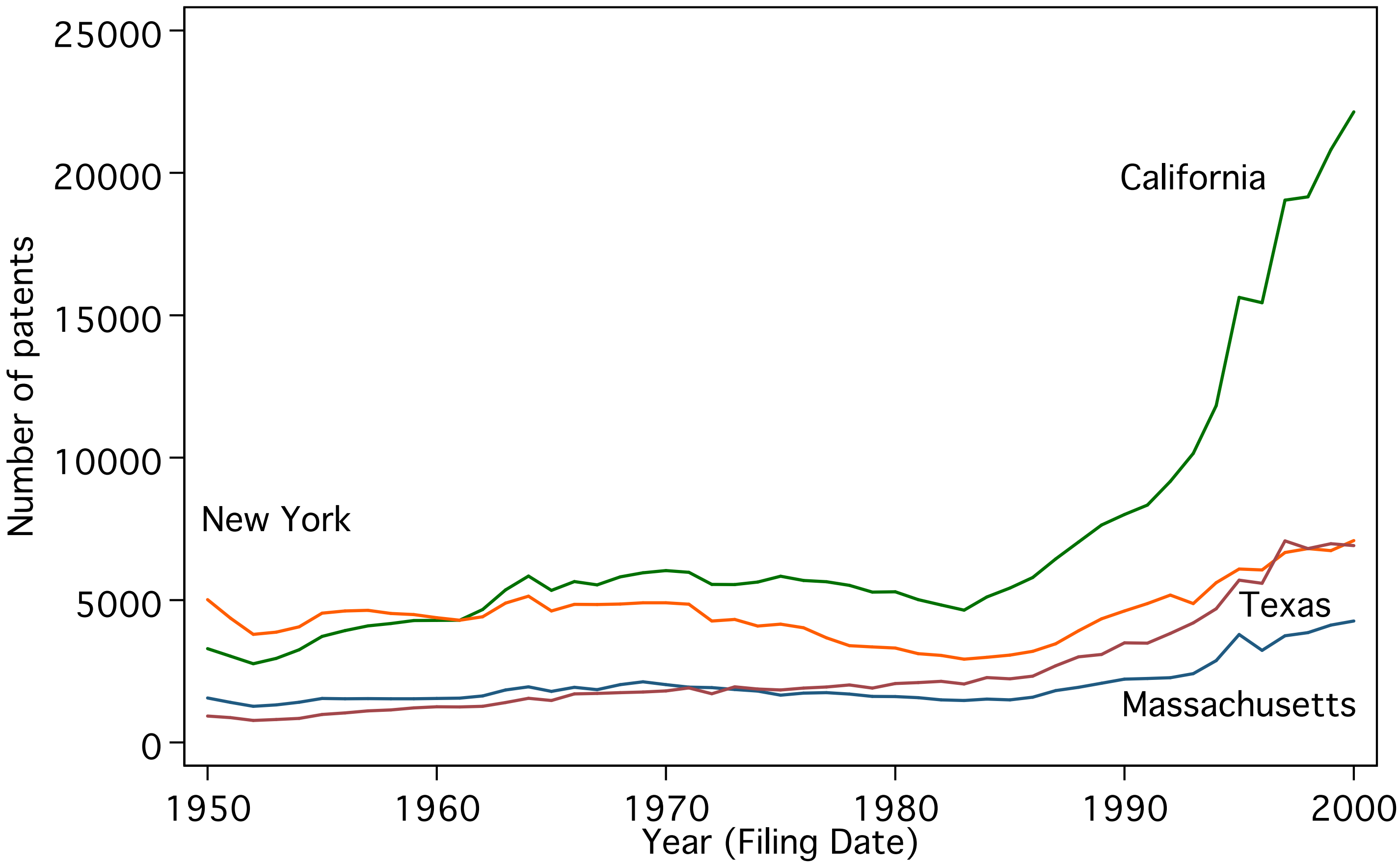
Sources:
Education, age, nativity: U.S. Census Bureau, IPUMS decennial census microdata usa.ipums.org/usa/
Return to college: author's calculations using IPUMS for wage regressions.
Orchestras: League of American Orchestras membership directory 2006 - www.americanorchestras.org/utilities/orchestra_members.html

Figure 1: States with Most Patents 1950



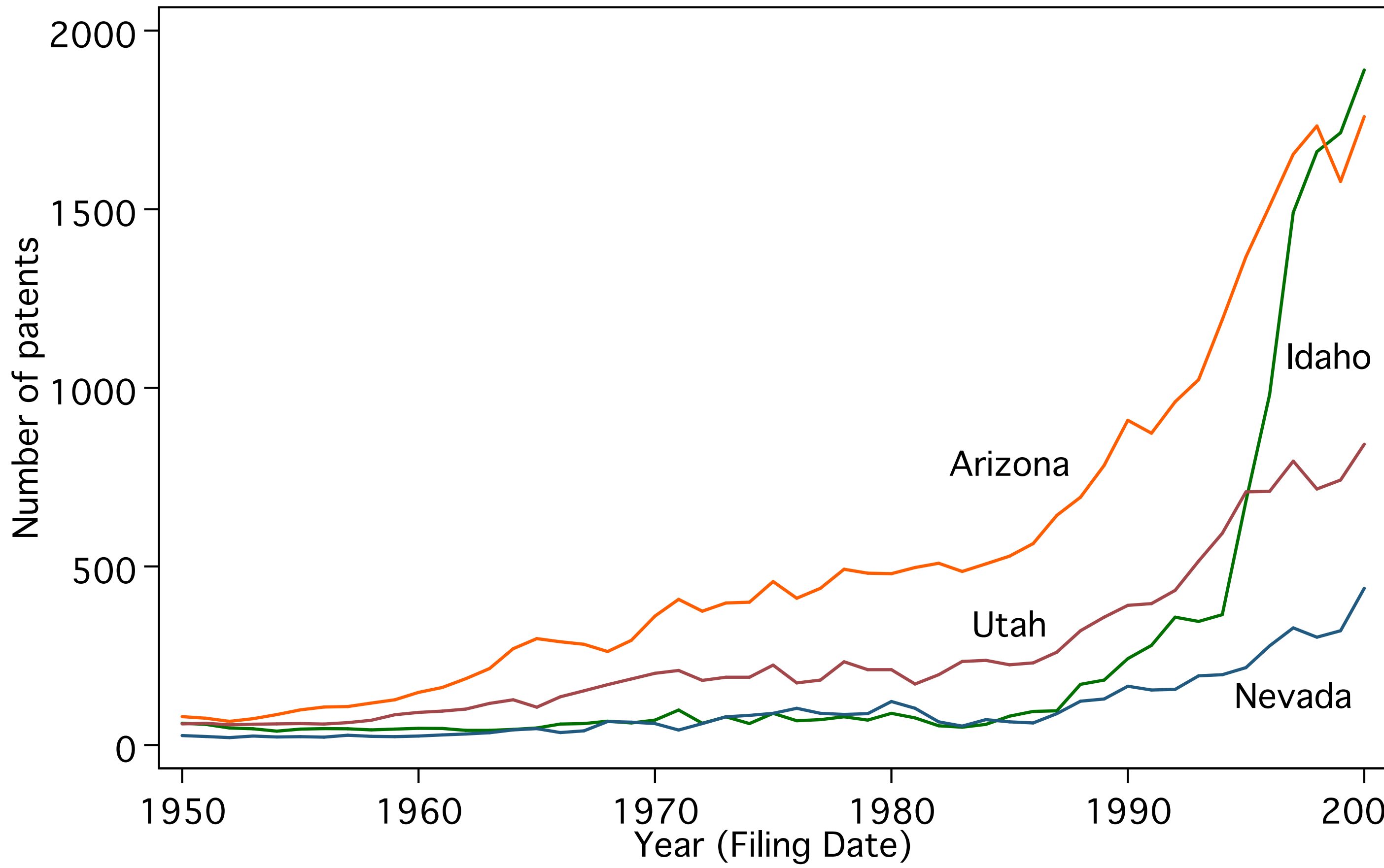
Source: USPTO and author's calculations

Figure 2: States with Most Patents 2000



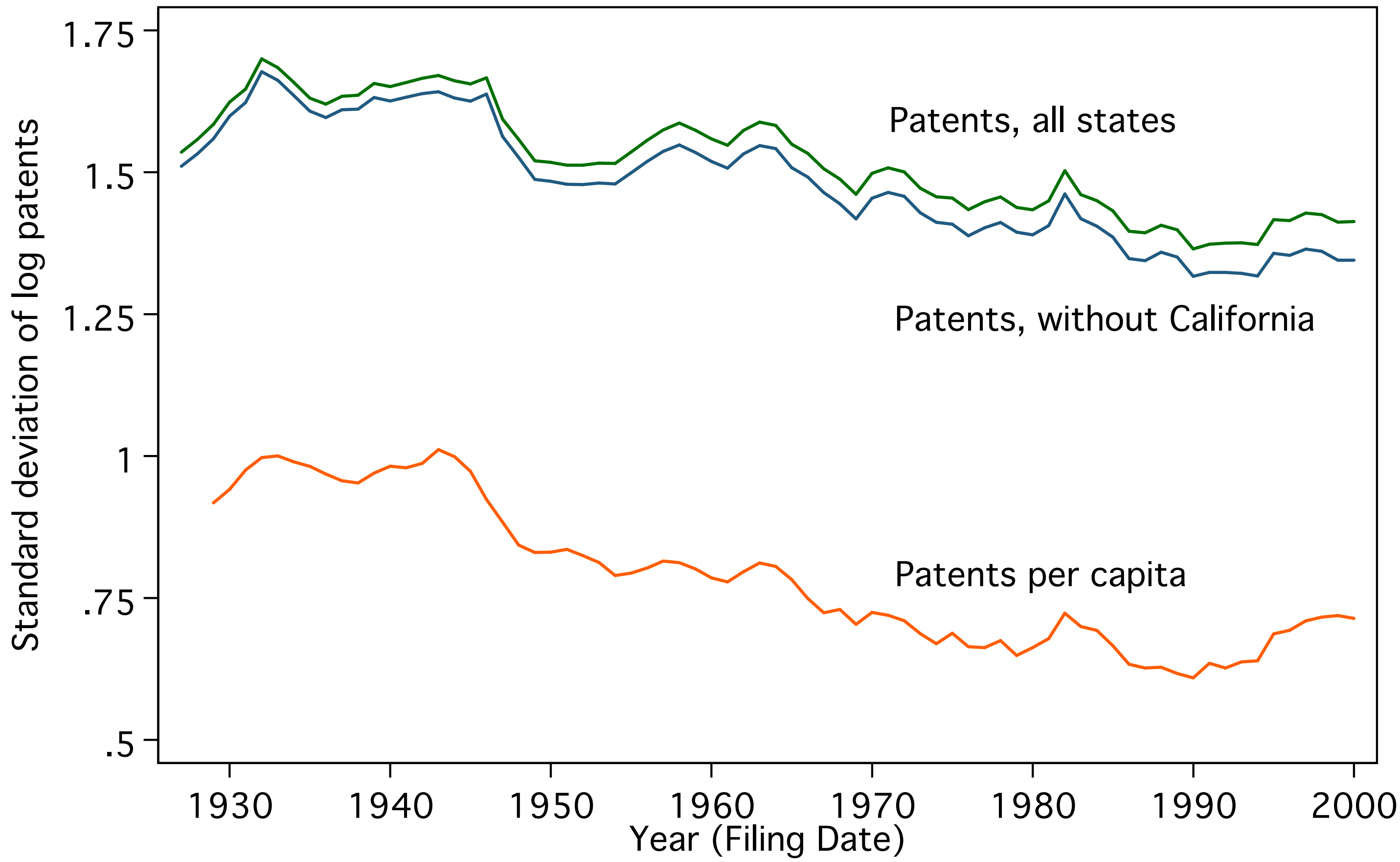
Source: USPTO and author's calculations

Figure 3: States with Fastest Patent Growth 1950-2000



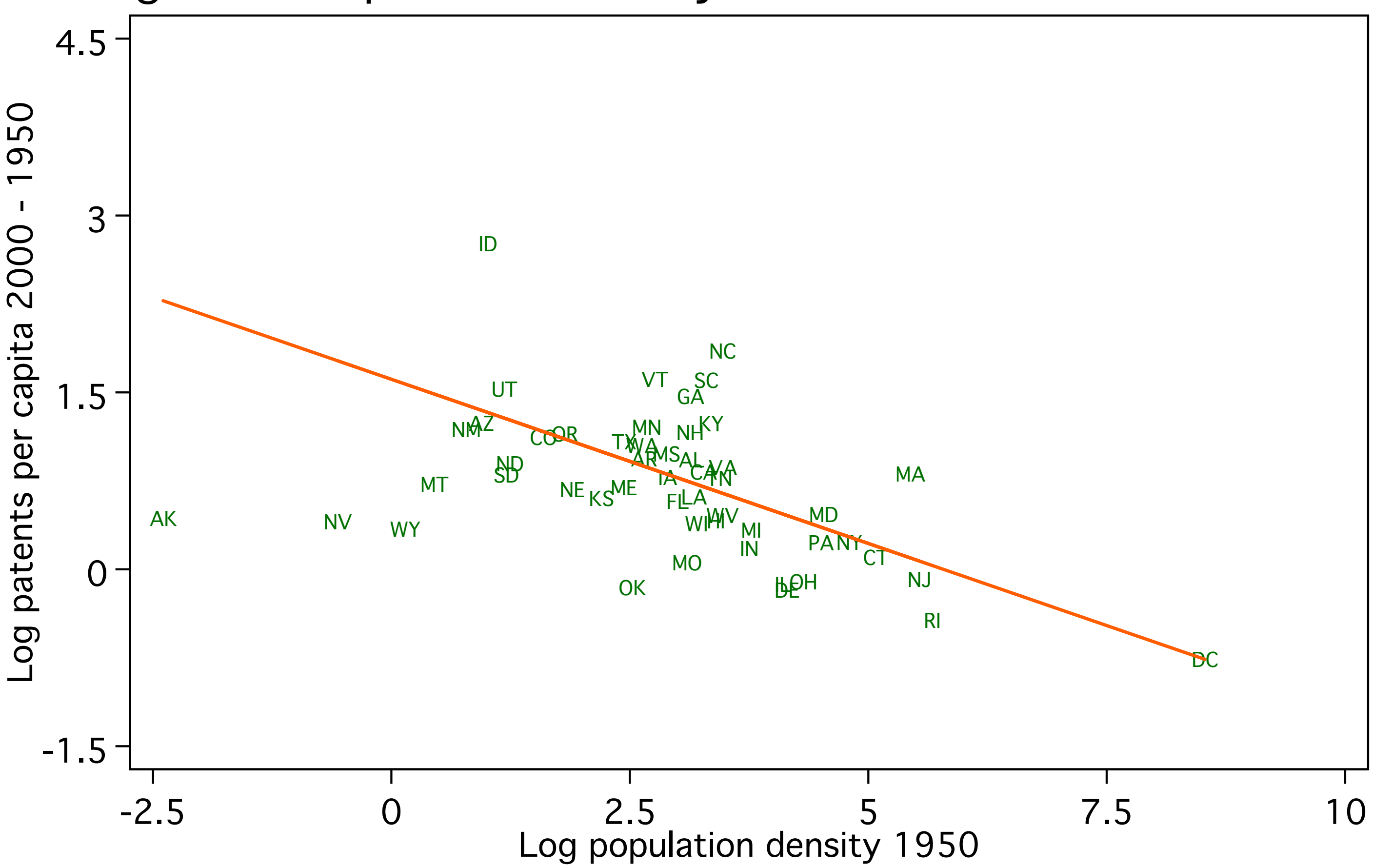
Source: USPTO and author's calculations

Figure 4: Convergence in Patenting Across States



Source: USPTO, BEA and author's calculations

Figure 5: Population Density as Predictor of Patent Growth



Source: USPTO and author's calculations