Immigrants and Innovation in the United States
Harriet Duleep and David Jaeger

Introduction

What are the benefits of immigration? A small body of literature has examined a variety of benefits of immigration to the U.S.: how they affect GDP (Borjas 1999, Hanson 2005), how they affect the efficiency of the labor market (Borjas 2001), whether they are a net fiscal gain or drain (a variety of studies in Smith and Edmonston, 1997, 1998), and how much they lower price levels (Cortes 2005). To our knowledge no studies have examined how immigrants affect the level and character of entrepreneurial activity in the U.S.

This paper examines how the presence of immigrants, apart from the entrepreneurial activity of the immigrants themselves, facilities innovation and entrepreneurial activity. If, as we hypothesize, immigrants' ability to learn new skills and methods at a lower opportunity cost than observationally equivalent natives fosters and encourages innovation and entrepreneurship (measured by self-employment, net business formation, and patent applications), immigrants are potentially a greater benefit to the U.S. economy than previous research might suggest.

Theoretical Model

We first posit a theoretical model that predicts that immigrants, particularly those possessing skills that are not immediately transferable to the U.S. labor market, facilitate innovation and entrepreneurship by being willing and able to invest in new skills.

At the heart of our theoretical prediction that immigration facilities entrepreneurship is the insight that human capital not immediately valued in the U.S. labor market may still be useful for learning new skills and therefore facilitate innovation. Persons who have learned one set of skills—even if those skills are not valued in U.S. labor market—have advantages in learning a new skill set. Cognitive psychologists refer to this phenomenon as "transfer." Because immigrants face a lower opportunity cost of investing in new skills or methods, this "transfer" of source-specific skills to the U.S. may lead immigrants to be more innovative and flexible in their human capital investments than observationally-equivalent natives. Areas and time periods with large numbers of immigrants (even if they are not self-employed) may, in fact, prove to be areas and time periods in which innovation is easier to accomplish.

We use these behavioral science perspectives to develop a model to test the impact of the flexible application of immigrant human capital in the US labor market. In deciding whether to develop a new product or service, potential entrepreneurs examine the costs and returns of pursuing such an activity. Returns are affected by the potential demand for a new product or service. In addition to capital outlays, a crucial cost of any new venture, particularly an innovative one, is training the workforce that will create the new product or service.

New businesses require people who are willing to acquire new human capital. Moreover, because immigrants have lower opportunity costs of learning new skills (because not all of their human capital is immediately transferable to the U.S.), employers may find it less costly to innovate in areas and time periods with large numbers of

immigrants. Immigrants of a given level of human capital will be less costly to train than natives. An entrepreneur in an area or time period with immigrants will have a relative advantage in launching an innovation. Conditional on the level of human capital, the advantage will be greater the lower the skill transferability of immigrants and the more innovative the product.

Data to Test the Model

Our model predicts that immigration yields higher levels of innovation as evidenced through increased entrepreneurial activity and business formation. Crucial to the analysis is defining entrepreneurial activity. We utilize several definitions of entrepreneurial activity for which we have acquired data over time. Our analysis is at the state level because data are not defined consistently at smaller geographic levels over the time periods we examine (roughly 50 years from the mid 1950s to 2005).

While we know of no publicly available measure of business formation that is collected in all states in a consistent manner, a reasonable proxy can be calculated from establishment counts available over time in both the Economic Census and County Business Patterns. Each data source identifies total counts of business establishments in a state. The Economic Census was conducted in 1954, 1958, 1963, and then every five years since 1967. County Business Patterns is an annual survey conducted each year since 1977. Differencing establishment counts over two periods of time provides a proxy for business formation activity in a state (although it captures both birth of new establishments and death of older establishments).

A second measure of business formation is self-employment data available from US decennial Censuses. In order to insure comparability over time, we utilize Census data from the Integrated Public Use Microdata Samples (IPUMS). For each decennial Census since 1950, we are able to determine the total number of individuals in non-farm self-employment as well as total income from these activities. It is important to note that our measure here is of overall entrepreneurial activity, not just that of immigrants.

A third source of data traditionally utilized to measure innovation is patent production. While our theoretical approach defines innovation as the marketing of new products and services, regardless of their patenting potential, the production of patents provides evidence of the broad innovative capacity of a locality. We employ the NBER Patent Citations Data File (Hall, Jaffe, and Trajteberg 2001), which has information on almost 3 million patents granted in the U.S. between 1963 and 1999. Most important for our research, the database identifies the geographic location of the inventor. Patents per capita and change in patenting provide another possible proxy for business formation and innovation.

Measuring immigration concentrations is also crucial to the analysis. We construct measures of foreign-born residents from two data sources: IPUMS decennial Census data (from 1940 to 2000) and the Current Population Survey (state-level data available annually from 1967). We also use measures of immigrant inflows available in data from the Immigration and Naturalization Service. These data have the distinct advantage of identifying immigrants who enter the U.S. under different visa categories. In all of these datasets we identify the country of origin of the immigrant as well as some measure of their skills (education or occupation).

Our analyses also require some measure of the demand for the businesses and the services and products they supply. Unemployment data provide a simple proxy for this demand. Data on unemployment at the state level are available from the U.S. Bureau of Labor Statistics (calculated primarily from CPS data) for every year from 1948 onwards. State-level population demographic characteristics of the native-born population, like educational attainment levels and age distribution, may also affect a state's business climate. Further, these same demographic traits of the immigrant population may change how they affect business formation. Data on the age and educational attainment of both the native-born and immigrant population are available both from the IPUMS decennial Census data and the Current Population Survey.

Empirical Specification of the Model

The empirical model utilizes regression analysis to assess how immigration and changes in immigration rates independently affect innovation and business formation across states over time, controlling for other relevant factors that may affect business growth. Immigrants are clustered in only six states with the largest percentage in California followed by New York. Within those states, immigrants are concentrated in the largest cities or metropolitan areas. It therefore seems likely that estimates of the effect of immigration based only on cross-sectional data would be sensitive to economic circumstances in any of the six states with heavy immigration. For instance, a downturn or upturn in California's economy (not captured by the explanatory variables in the regression) might be captured by the immigration variable.

The basic form of our empirical specification is

$$BD_{tj} = \alpha + \beta I_{tj} + \gamma U_{tj} + Z_{tj}\Phi + \lambda_t + \gamma_j + \varepsilon_{tj} \tag{1}$$

where $j = 1 \ldots n$ and $t = 1 \ldots T$, BD_{tj} is business development in time t and area j, I_{tj} is immigration share of workforce at time t in area j, U_{tj} is unemployment (as a measure of demand) in time t and area j, Z_{tj} is a vector of variables measuring education and age mix of U.S.-born and immigrant working-age population in time t and area j and, λ_t and γ_j are time and state fixed effects, respectively. We estimate this model for a variety of time frames, depending on the data used and the definition of business development that we can employ using those data, as outlined above.

It is also possible to estimate this fixed-effects model as a first-difference model. In the fixed-effects model and the first-difference model, however, endogeneity bias is still potentially a problem if immigrant concentrations are correlated with the time-varying error term. This might happen, for example, if immigrants are drawn to areas with a high degree of entrepreneurship. One solution to this problem is to use instrumental variables estimation, with lagged values of I_{tj} as the instrument. Another potential solution is to use plausibly exogenous changes in immigrant locations as "natural experiments," e.g. the influx of refugees after World War II, Cuban migration following Fidel Castro, the large changes associated with the elimination of the quota system after 1964, or the influx of Southeast Asian refugees after the Vietnam War. All of these changes are plausibly uncorrelated with entrepreneurial activity and thus provide an exogenous change in immigrant concentrations across time and space.

One final source of variation that we explore is how immigrants from different countries affect entrepreneurial activity. Our theory predicts that because the skills of immigrants from less-economically developed countries are less likely to be directly transferable to the U.S. they will have a greater impact on entrepreneurial activity. All of our data sources permit us to stratify the immigrant population by country of origin. An exploration of differences in both the magnitude and significance of the impact of immigrants from more- and less-developed countries on business development provides a means to both address the issue of simultaneity and provide additional verification of the theory. We also control for the entrepreneurial behavior of immigrants by adding information on the proportion of immigrants who are self-employed or the proportion of self-employment income earned by immigrants. Finding positive and significant coefficients on our immigrant variables while controlling for immigrant entrepreneurial behavior would add further confirmation of our unique theoretical approach.

Summary

Using historical and cross-sectional data, this analysis explores the impact of immigration on state-level business formation and entrepreneurial activity. Our theory offers a unique perspective on the contributions of immigrants to economic development beyond traditional perspectives that focus on low-cost immigrant labor or immigrant entrepreneurship.

References

Borjas, George J. (1999) "The Economic Analysis of Immigration," in Orley C. Ashenfelter and David Card, eds*., Handbook of Labor Economics*, Amsterdam: North- Holland, 1697-1760.

Borjas, George J. (2001) "Does Immigration Grease the Wheels of the Labor Market," *Brookings Papers on Economic Activity*, 1:2001, 1-51.

Cortes, Patricia (2005) "The Effect of Low-skilled Immigration on US Prices: Evidence from CPI Data," unpublished mimeo, MIT, November.

Hall Bronwyn H., Adam B. Jaffe, and Manuel Trajteberg (2001) "The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools," NBER Working Paper 8498, November.

Hanson, Gordon (2004) "Immigration Policy," unpublished mimeo, University of California-San Diego, July.

Smith, James P. and Barry Edmonston, eds. (1997) *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration.* Washington, DC: National Academies Press.

Smith, James P. and Barry Edmonston, eds. (1998) *The Immigration Debate: Studies on the Economic, Demographic, and Fiscal Effects of Immigration.* Washington, DC: National Academies Press.