Appendix to
Ethnic Scientific Communities and International Technology Diffusion

William R. Kerr
Harvard Business School and MIT

February 2007

1 Introduction

This technical appendix collects theoretical analyses, data construction notes, and empirical results of Kerr (2007a) that were included in working papers but are omitted from the published paper. Section headings (but not necessarily section numbering) correspond to the published paper. Comments are appreciated and can be sent to wkerr@hbs.edu. Please reference the main paper when citing this material.

2 Theoretical Framework

2.1 Agriculture to Manufacturing Sector Reallocation

The paper outlines a simple technology transfer model between a frontier country and a following nation. The steady-state characterization of the follower's economy builds on the assumption of full employment in the manufacturing and research sectors. While the estimating equation (6 in main paper) relates the follower's output to its ethnic research presence in the leader, the same elasticity β would hold for labor productivity specifications. With full employment, output gains can only come through labor productivity enhancements. Many developing economies have large agricultural sectors, however, and the migration from agriculture to manufacturing is important for characterizing economic development (e.g., Harris and Todaro 1970).

This section incorporates into the basic model an agricultural sector in the follower. The framework highlights how technology transfer induces a different response when sector reallocation is possible. Specifically, as more technologies are transferred from the leader to the follower, labor shifts from agriculture to the manufacturing and research sectors. After a sufficient number of frontier innovations have been imitated, the follower's economy transitions to full employment in the manufacturing and research sectors. Thus, the steady-state of the expanded economy is the same as the basic framework described above; numerical simulations of the transition path, however, offer additional guidance for the empirical exercises this study undertakes.

The agricultural sector for the follower is characterized by a decreasing returns to scale technology that employs only labor L_A,

$$Y_A = BL_A - \frac{1}{2}L_A^2. \tag{1}$$

B is a common agricultural productivity parameter, and the final goods from agriculture and manufacturing are identical ($Y = Y_A + Y_M$). Labor is again free to move across sectors, and the proportion of the labor force allocated to each of the three sectors along the development path can be related to the follower's technology stock. The follower is assumed to possess only imitated technologies ($M = N$).

First, the marginal products of labor for agriculture and manufacturing are $B - L_A$ and $(1-\alpha)A^{1/(1-\alpha)}\alpha^{2\alpha/(1-\alpha)}M$, respectively. Wage equality between these two sectors relates the size of the agricultural workforce to the number of imitated technologies,

$$L_A = \max[B - (1-\alpha)A^{1/(1-\alpha)}\alpha^{2\alpha/(1-\alpha)}M, 0]. \tag{2}$$

Thus, growth in the follower's technology stock lowers agricultural employment until the economy reaches a transition point with full employment in the research and manufacturing sectors. This transition occurs when $M > (1-\alpha)^{-1}A^{-1/(1-\alpha)}\alpha^{-2\alpha/(1-\alpha)}B$. If this condition is satisfied in the current period, it will hold in all future periods as the wages of the manufacturing and research sectors continue to grow with further technological advancement.[1] Likewise, the size of the manufacturing labor force can be related to the follower's existing technology stock and the frontier technology stock through the wage equality of the manufacturing and research sectors and the interest rate $r = \rho$,

$$L_M = \frac{M}{\tilde{I}}\Psi^{-1}\left[\frac{M}{\tilde{I}}\right](\tilde{H}^F)^{-\beta}\frac{\rho}{\alpha}. \tag{3}$$

With L_A and L_M determined, the number of researchers follows from the labor endowment.[2]

To characterize the transition path, a numerical solution for the steady-state without an agricultural sector is first developed. The labor forces of the two economies are taken to be of size 0.20. For parameter values of $\alpha = 1/3$ and $\rho = 0.05$, the allocation of labor to manufacturing and research is 0.15 (75%) and 0.05 (25%), respectively. As the growth rate of inventions in the leader is equal to the size of its research labor force (i.e., $(\partial\tilde{I}/\partial t)/\tilde{I} = \tilde{L}_R$), the set of available frontier technologies grows at a rate of 5%; the same growth rate is in turn found for the follower's imitated technology stock and the manufacturing outputs of the two countries. Specifying $A = 1$, the steady-state value of new inventions or imitations in both economies is given a numerical value of $V = 0.22$.

Examining more closely the technology transfer mechanism, the size of the follower's ethnic research population living in the leader is modeled as 0.001 (or 2% of the frontier researcher total). Taking an estimate of $\beta = 0.3$ from the empirical exercises in Section 4 and assuming a depreciation on human capital of $\delta = 0.15$, the steady-state human-capital stock is given a numerical value of $\tilde{H}^C = 0.0067$. Finally, a functional form for $\Psi[M/\tilde{I}]$ must be specified to estimate the share of leader technologies imitated by the follower. The form $\Psi[M/\tilde{I}] = 0.5\cdot(1 - M/\tilde{I})$ retains the properties of $\Psi' < 0$ and $\Psi[1] = 0$ and yields a steady-state imitation share $M/\tilde{I} = 0.10$.

Turning to the transition path simulations to this steady-state, the solid line in Appendix Figure 1 describes the evolution of the follower's economy from the initial conditions of 90%

[1] The agricultural production function (1) bounds the marginal productivity of labor from above at B. This function does not satisfy the Inada conditions. An agricultural sector with a constant returns to scale production function employing land and labor yields an ever shrinking agricultural sector.

[2] Even if the follower's human-capital stock for frontier technologies is stable during the transition period, the relative proportion of labor devoted to manufacturing versus research is not at its steady-state level due to adjusting fraction of imitated technologies $M/\tilde{I}$.

employment in agriculture and 1% of US technologies imitated.[3] The size of the agricultural sector corresponds to an initial stock of imitated technologies $M(0)$ in the follower, while the gap to the frontier determines the technologies $\tilde{I}(0)$ in the leader; the parameter $B = 0.2$ is also specified. The leader, assumed to be in steady-state growth, evolves exogenously with a 5% growth rate in its technology stock. From (2), (3), and the labor endowment, the evolution of the follower's economy is subsequently characterized.

The leftmost panel describes the allocation of labor along the transition path. For the baseline simulation, the follower's human-capital stock with respect to frontier technologies remains at its steady-state level of 0.0067. Initially, limited labor is devoted to manufacturing or research. In this general equilibrium, the small market size of final-goods producers depresses the value of new innovations and the researchers employed, even though the large gap to the frontier makes it very productive to imitate new intermediate products. Likewise, the labor demand of final-goods producers is limited due to the small technology stock in the follower.

As the number of technologies steadily expands, however, the agricultural sector shrinks and more labor is allocated to both manufacturing and research. This industrialization sustains itself as the growth in market size increases the value of new innovations, while the larger technology base increases the labor demand of final-goods manufacturers. The sector reallocation quickens as the economy approaches the transition point ($t = 16$). Around this transition, the researcher share of follower's labor force reaches its peak, before gradually declining to its steady-state value of 25%. The manufacturing labor share also surges around the transition, and continues to grow to its steady-state share of 75%.

The middle panel presents several growth rates evident in the follower during the transition. Initial growth is slow due to the inertia of the large agricultural sector. Around the transition point, however, growth in the follower's imitated technology stock surges due to the extensive labor resources devoted to research and the still sizeable gap to the leader's frontier. This high rate of technology adoption translates into higher growth in both manufacturing output and labor productivity. The manufacturing output growth is not due solely to labor productivity gains, however, as the growth in employment contributes approximately the same amount. After the transition, the growth rates decline to their steady-state rates of 5%.

Finally, the rightmost panel exhibits several levels with respect to the frontier. As evident in research labor share, the linear preferences of consumers affords a substantial investment in the imitation of new technologies around the transition in return for higher future consumption. During this convergence period, the follower's technology gap to the frontier is substantially narrowed. In the steady-state, the fraction of frontier technologies imitated by the follower translates directly into the steady-state fraction of frontier manufacturing output achieved, and the transition path dynamics take the same shape. Note, however, that the follower's total

[3] Full employment in agriculture is an unstable equilibrium if the follower's human capital to frontier technologies exists. Once some labor is devoted to manufacturing and research, the follower's economy will eventually transition completely out of agriculture if access to frontier inventions is maintained.

output level (including agriculture) relative to the frontier declines slightly during the transition period due to the investment in imitation. Except in the immediate vicinity of the transition point, the follower's output does not fall but instead fails to maintain pace with the frontier economy in steady-state growth. After this investment period, however, the share rises sharply to a long-run level equal to the manufacturing output share.

From this baseline, the dotted line in Appendix Figure 1 plots a second transition path for an exogenous increase in the number of researchers of the follower's ethnicity living in the leader from 0.001 to 0.004 on date $t = 3$.[4] While these simulations are meant to be illustrative, the fourfold rise from 2% to 8% of the US research community is roughly in line with the growing Chinese research contribution in several high-tech industries for the period studied in the paper's empirical analysis. As the top left panel shows, this exogenous increase does not immediately translate into a fourfold increase in the follower's human-capital stock with respect to frontier technologies. The human-capital stock instead grows over time with the higher rate of codified and tacit knowledge gain in each period following the US ethnic researcher growth.

As the boost in technology transfer is realized, however, the transition from agriculture proceeds at a more rapid pace. The growth rate of manufacturing output spikes upward due to both higher growth in imitated technologies and more labor reallocation. An economy without an agricultural sector would only experience output growth due to labor productivity gains. In the new steady-state, the fourfold increase in the follower's human-capital stock results in an approximate 40% levels gain in imitated technologies and output; the percentage of frontier technologies imitated is also higher. The follower's growth rate and allocation of labor, though, are the same as in the simulation without the exogenous increase in scientific integration.

In summary, technology transfer to economies with large agricultural sectors can increase manufacturing output through both labor productivity gains, as in the steady-state scenario presented in Section 2 of the main paper, and through employment growth. In this particular framework, the productivity gains and employment gains are of roughly similar magnitude. In alternative models, however, output growth would come only through labor reallocation. An example is a specification with constant outside wages and a fixed stock of physical capital. As the technology transfer increases the marginal product of labor, producers hire more labor to bring the marginal product of labor back down to the external wage. The agricultural sector's production function (1) instead allows the marginal product of labor in agriculture to increase in step with the manufacturing sector's wage.

2.2 Additional Notes

- As in most endogenous growth frameworks, the perpetual monopoly rights assumption can be relaxed in both nations. This would allow for imitation among research firms within either the leader or the follower.

[4] The simulations abstract from any growth in the overall size of the leader's labor force due to this inflow.

- The model assumes competitive markets in both the leader and the follower; this may be the framework's most questionable assumption given the wide range of developing, emerging, and industrialized economies to which it is applied.

3 Ethnic Patenting and International Citations Analysis

3.1 Ethnic Patenting Records

Kerr (2007c) documents the name-matching algorithms, lists frequent ethnic names, and provides extensive descriptive statistics.

3.2 International Patent Citation Analysis

As outlined in the main text, citation counts are developed by cells that contain four dimensions: 1) the ethnicity of the citing foreign inventor, 2) the ethnicity of the cited US inventor, 3) the technology class of the citing foreign inventor, and 4) the technology class of the cited US inventor. The latter two dimensions are necessary for isolating ethnicity's role since patents cite other patents within their technology field far more frequently than those outside of their field. If ethnicities concentrate in different industries in the US and abroad (e.g., greater Chinese research in computer industries), measured ethnic flows could be merely capturing that technologies build upon prior art in their own discipline.

Almost 100,000 cells are formed with this data organization, and many cells contain zero values. The zero values are due to both the small sizes of some ethnicities (e.g., Vietnamese inventors outside of the US) and that researchers in a given field simply do not cite the universe of technologies in their work. Count data containing zero values can be appropriately handled with a Negative Binomial model.[5] The counts are regressed on an indicator variable for whether the citing foreign ethnicity and cited US ethnicity are the same, as well as vectors of fixed effects for each of the four dimensions on which cells are formed. These fixed effects remove basic levels differences between the series (e.g., English in the US receiving uniformly more citations, Vietnamese researchers abroad making uniformly fewer inventions and citations). An indicator variable is also included for whether the cited and citing technology categories are the same.

The coefficient on the indicator variable for same-ethnicity is transformed into an incidence rate ratio that gives the higher rate of citations within an ethnic group. Panel A of Appendix Table 1 tabulates the incident rate ratios graphed in Figure 2 of the main paper. Again, these coefficients are compared to a baseline value of one, the level where own-ethnicity citations have the same frequency as citations of other ethnicities. As the main text relates, the results suggest a moderate effect that own-ethnicity citations are 50% higher than citations to other ethnicities once the basic levels and industry effects are removed. Moreover, common ethnicity appears

[5] Wooldridge (Ch. 19, 2002) describes the statistical properties of this empirical framework. The data reject the Poisson model in favor of the Negative Binomial model.

most important for international technology diffusion in the first few years after an invention, peaking in a citation lag of four to five years.

These estimations employ fixed effects and an indicator variable for common technology fields. Due to computation restrictions, these technology fields are defined at the two-digit subcategory level of the USPTO framework (36 groupings).[6] These patterns results are easily extended, however, to the three-digit USPTO patent classes (about 450 groupings) when dividing the sample into four groups based upon broad technology categories.[7] These results are shown in Panels B-E of Appendix Table 1. These more narrow groupings suggest a 20%-30% ethnic differential. While the lag structures continue to emphasize the first five years of the diffusion process (Panel F), some interesting heterogeneity is present. Ongoing research is attempting to characterize these differences.

Nevertheless, Thompson and Fox-Kean (2005, hereafter TFK) criticize even the three-digit USPTO categories as being too broad to control effectively for technology specialization. TFK instead employ the sub-class level where over 150,000 divisions exist, while the response by Henderson et al. (2005) emphasizes the liabilities of the sub-class level vis-a-vis the three-digit classifications.[8] This debate raises questions about best practices for controlling for technology fields in citation-based analyses, although it is encouraging for the present study to note that the TFK criticism mostly applies to intra-national citations rather than inter-national results.

As an alternative, Thompson (2006) proposes comparing inventor-added citations to those added by the USPTO examiner. This valuable distinction for citations is only made for patents granted after 2001. The concept here is that the examiner-added citations, all made from the USPTO office in Virginia, can serve as an effective control for examining geographic localization (as in Thompson) or ethnic localization (here) as the examiner makes his or her citation additions independently. The advantage of this technique is that it avoids using the USPTO technology fields entirely. Thompson (2006) further describes the identification strategy and its potential limitations.[9]

This study uses the Thompson (2006) technique as a robustness check on the main findings, although this requires moving to a new time period from the one studied in the main paper. Thompson (2006) constructs a dataset of all patents granted by the USPTO in January 2003 and their citations; this dataset contains the additional distinction of whether citations are made by the inventor or added by the examiner. Thompson graciously provided access to this data,

[6]The 93,312 observations are from the dual crossing of 36 citing and cited technology subcategories and 9 citing and cited ethnicities, with English inventors abroad excluded (36x36x9x8). This latter exclusion is made for conceptual reasons but not very important.

[7]The four divisions correspond to the highest level USPTO categories, with Computers and Communications, Drugs and Medical, and Electrical and Electronic combined.

[8]Agrawal et al. (2006) further refine the TFK procedure.

[9]The geographic control is based upon the Alexandria, VA, location of the USPTO office. It is not clear that this would directly translate to an ethnicity context. To confirm that the ethnicity of the examiner would not bias the control group, the name-matching algorithm is also applied to examiner names. Estimations did not find that examiners are more likely to add citations of their own-ethnic groups compared to inventor-added citations.

including inventor names, and this study assigned ethnicities to these inventors in the same manner as the primary analysis. To match this paper's conceptual framework, Thompson's sample is restricted to include only USPTO patents filed from outside of the US and their citations of past patents that were filed inside of the US (4835 count). Both inventor names and ethnic assignments are successfully collected for 2973 observations.

A dichotomous indicator variable for same ethnicity is again created. It takes on a value of one if the cited and citing inventor are of the same ethnicity and zero otherwise. The sample means are themselves informative. Including English inventors abroad, inventor-added citations have an own-ethnicity mean of 0.245 versus just 0.144 for examiner-added citations. That is, about 25% of the US-based citations made by foreign inventors are within the inventors' ethnic groups, versus just 14% for those made by independent examiners. This differential is larger in relative terms when excluding English inventors abroad: 0.028 versus 0.009.[10]

Following Thompson (2006), these sample descriptive statistics are refined through conditional logit estimations in Appendix Table 2. The odds ratios are presented with z-scores in parentheses. The estimations include citing patent fixed effects, and the 889 observations are the effective number of observations after removing those citations fully accounted for by the fixed effects. The first column only includes the indicator variable for inventor-added citations. The results suggest a 56% gain that is statistically significant at the 10% level. While developed in quite different ways, this point estimate is quite similar to the Negative Binomial results presented earlier. Column 2 next adds a second indicator for a non-institutional citation and the time lag in years from the date of cited invention to the citing invention. While the same-ethnicity point estimate remains economically significant at 43%, the coefficient is no longer statistically different from one, the level where no ethnicity effect is found.

The third column, however, shows that this decline comes in the citations of patents that are over ten years old. This expanded estimation introduces a second indicator variable for citations of patents ten years or older and an interaction of this age indicator with the inventor-added citations dummy. The results resemble the time path reported in Figure 2 of the main text. The own-ethnicity effect is quite strong (76%) for citations of patents invented within the previous ten years, but no same-ethnicity effect is present for citations of patents ten years or older. These results again suggest a stronger within-ethnicity effect at the start of the diffusion process, but there are unfortunately too few observations to estimate year-by-year time lags like those shown in Appendix Table 1.

This study views both the earlier Negative Binomial estimations and the Thompson (2006) framework as but a first pass at understanding the role of ethnicity in international inventor-to-inventor knowledge diffusion. The results suggest an important ethnic channel for international technology transfer and thereby motivate the primary estimations of macroeconomic gains

[10] These latter means may appear low, but recall 1) over three-quarters of inventors in the US are English, 2) almost half of foreign inventors are Japanese and the Japanese community in the US is very small by comparison, and 3) the within-industry dimension is more important than the within-ethnicity dimension.

through greater scientific integration. It is hoped that future research will refine our understanding of these important issues, especially as additional data emerges and the best practices for accounting for technology specialization are refined.

4 Output and Productivity Analysis

4.1 Foreign Manufacturing Data

The benefit of knowledge integration for foreign development is evaluated through the Industrial Statistics Database of the United Nations Industrial Development Organization (UNIDO). To complement Table 2 of the main paper, Appendix Table 3 provides additional descriptive statistics for the primary panel constructed. Some additional notes regarding the UNIDO data:

- The UNIDO collects data at the three-digit and four-digit industry levels of the International Standard Industrial Classification (ISIC). The paper focuses on the three-digit aggregation, but the four-digit data delivers similar results. While sacrificing industries (28 versus 80), the three-digit dataset contains more countries (43 versus 20), better coverage of Chinese economies, and more capital data.

- The USPTO issues patents by technology categories rather than by industries. Combining the work of Johnson (1999) and Silverman (1999), concordances are developed between the USPTO classification scheme and the three-digit industries in which new inventions are manufactured or used. The main estimations focus on industry-of-use, affording a composite view of the technological opportunity developed for an industry. Studies of advanced economies find accounting for these inter-industry R&D flows important (e.g., Scherer 1984). Keller (2002) reports inter-industry R&D flows aid productivity growth significantly within OECD countries, equal to half or more of the own-industry development. Estimations with manufacturing industries support the using-industry specifications.

- The UNIDO dataset is inappropriate for studies of industry creation or destruction due to its unbalanced panel and industry aggregation. Recognizing this limitation and in order to enhance the quality of estimations, country-industry observations must maintain ten employees and one US ethnic patent per annum. These minimums exclude poor quality data, but raising or removing these hurdles does not significantly affect the findings.

- Specifications employing alternative UNIDO data on industry value-added and establishments mirror the output and employment results presented. The 1985-1997 period balances data inclusion with maintaining a consistent sample, as data for earlier or later years are quite limited. Similar outcomes are evident if all 1980-2000 data are employed or if the sample is restricted to a 1985-1997 balanced panel of continually surveyed countries and industries.

- Some country-to-ethnicity mappings are debatable (e.g., placing Spain and Portugal with European rather than Hispanic, including the Scandinavian countries in European), as is the inclusion of communist countries. The results are robust to reclassifying or dropping these marginal cases, as appropriate, on a case-by-case basis.

4.2 Output and Productivity Estimation Framework

- The substantial increase in the number of patents granted by the USPTO over the last two decades is difficult to interpret. See, for example, Kortum and Lerner (2000), Kim and Marshcke (2004), Hall (2005), Jaffe and Lerner (2004), and Branstetter and Ogura (2005). Restricting the analysis to within-industry variation circumvents this issue.

- Industry-year effects extract industry-specific price movements (e.g., the rapid decline of computer prices). The UNIDO converts output data from foreign currencies to nominal US dollars using average yearly exchange rates (IMF International Financial Statistics Series rf).

4.3 Ethnic Patenting Estimator

- The five-year sums of recent ethnic patenting in the core estimating specifications give equal weight to each year. Regressions weighting the lagged community sizes by the coefficients from the international citation exercises yield similar results.

4.4 Basic Output and Productivity Regressions

- Appendix Tables 4-9 presents the levels specifications that parallel Tables 3-8.

4.5 Foreign Country Development Controls

- The physical-capital regressions are the best possible given the data. Sufficient capital data are only available for the countries noted in Appendix Table 3 and do not always cover the years listed in the UNIDO3 Panel column. Capital stocks are estimated using the perpetual inventory method with a depreciation rate of 15%. Initial stocks are developed using 1980 and 1981 investments, and subsequent investments are deflated using weighted deflators taken from the NBER-CES Manufacturing Productivity Database (Bartelsman and Gray 1996). Breaks in the capital series for Chile (1987, 1988), Macao (1987), Mexico (1992, 1993), Panama (1986), and Peru (1993) are bridged in the reported regressions; the results are robust to instead dropping the years after the breaks.

- A better test of the capital growth is with the labor productivity specification and using capital-labor ratios. This is undertaken in Appendix Table 10.

4.6 Sector Reallocation

- Table 2 of the paper lists the 1980 share of national employment in agriculture for each economy. Agricultural shares are from the United Nations Statistical Division and Sun et al. (2003).

- Appendix Tables 11 and 12 include the sample decompositions with the sector reallocation measurements.

5 Exogenous Changes from US Immigration Reforms

Section 5 opens with a simple thought experiment to illustrate the reverse causality concern. The follower's government temporarily subsidizes invention until Section 2's steady state conditions no longer hold. As $I > \tilde{I}\Psi[M/\tilde{I}](\tilde{H}^F)^\beta$, it is more profitable for researchers in the follower to invent rather than imitate; the follower's output growth and sector reallocation are now driven solely by domestic innovations. In the leader, researchers of the following country's ethnicity switch from inventing to imitating, as the latter is initially very easy (i.e., $\Psi[0]$ is high).

How the system evolves from this initial disturbance depends on the relative populations of researchers of the follower's ethnicity in the two countries. If the expatriate research group in leader is sufficiently small, they will continue imitating a large invention stock developed abroad forever, and the follower's human-capital stock with respect to the leader's inventions will decline to zero. The follower's domestic researchers will continue inventing, and the gap between the follower's researcher productivity for invention versus imitation will become entrenched. On the other hand, if the expatriate research group in leader is sufficiently large relative to the follower, the declining imitation productivity will require at least some expatriate researchers in the leader resume direct invention to maintain full employment. In this scenario, the initial reverse technology flows yield to either a sustained mixing strategy, with researchers of the follower's ethnicity in both countries inventing and imitating, or the leader resuming the inventing role.

5.1 The Immigration Act of 1990

- US immigrants are admitted through numerically restricted categories, governed by the quotas discussed in the main paper, and numerically unrestricted categories (e.g., immediate relatives of US citizens). The reduced-form estimator centers on the numerically restricted categories that admit 75% of ISEs, versus 43% of all immigrants. Jasso et al. (1998) outline US immigration policy and the 1990 Act; they further discuss behavioral responses to changes in quotas. ISE inflows through the unrestricted categories are stable in the years surrounding the 1990 reform.

- Immigration trends in Figure 3 of the main paper are developed from immigrant-level INS records. The graphed permanent residency admissions include ISEs already working in the US on temporary visas. The trends for "new arrival" ISE are very similar. Temporary visas can only be renewed once, so the total shift in ISE population should include workers gaining permanent residency. The analysis does not depend on this distinction. Science and engineering categories are defined as Engineers, Natural Scientists, and Mathematical and Computer Scientists; low-skilled categories are Administrative Support, Farming, Laborer, Precision Production and Repair, Service, and Sales occupations.

- Appendix Figure 2 documents trends in low-skilled immigration around the 1990 Act. While Chinese and Indian immigration are substantially higher than Hispanic immigration for science and engineering, the opposite is true for low-skilled immigration. Moreover, low-skilled immigration did not respond to the 1990 Act.

- Appendix Figure 3 documents the NSF trends in science and engineering Ph.D. graduates in the US by country of origin. On a logarithmic scale, Appendix Figure 3 exhibits a smooth trend for Mainland China from 1985-1991 with a marginal decrease in the growth rate thereafter. This graph was included in working papers to discuss the potential omitted variable biases involved in the reduced-form estimator.

- Appendix Table 13 provides the immigration preliminaries: the data for the immigration responses analysis below, the waiting list materials, and the unreported regressions discussed in Section 5.2 of the main paper.

5.2 Immigration Responses

The reduced-form strategy exploits differences in the extent to which countries were affected by the 1990 reform. It is inappropriate, however, to use the outcomes exhibited in Figures 3-4 of the main paper to determine treatment and control groups. A proper designation of the affected countries requires a more formal analysis of researcher immigration responses to the legislation change. Let $ISE\%_{ct_0}^{Adm}$ be the mean ISE arrivals from country c divided by an approximate country-level numerical limit for employment-based workers during the 1983-1990 pre-period. The theoretical numerical limit is taken to be the 20,000 country limit multiplied by the 20% worldwide allocation given to employment-based applications (i.e., 54,000/270,000).[11]

Define $POST_t$ as a indicator variable taking the value of zero from 1983-1990 and one for 1991 and after (i.e., the 1990 Act's effective date). Regressing annual ISE admissions ISE_{ct}^{Adm} on an interaction of $ISE\%_{ct_0}^{Adm}$ with $POST_t$ quantifies the immigration response of constrained countries,

$$ISE_{ct}^{Adm} = \alpha + \gamma ISE\%_{ct_0}^{Adm} \cdot POST_t + \phi_c + \eta_t + \epsilon_{ct}. \tag{4}$$

[11] The total employment immigration column in Appendix Table 13 demonstrates the theoretical limit works quite well. The scientific percentages are even larger than they initially seem since family members of employment-based admissions count towards the two quotas.

The main effect for $ISE\%_{ct_0}^{Adm}$ is absorbed by the country fixed effects ϕ_c, along with levels differences between nations in US immigration. The year effects η_t remove aggregate changes in US permanent residency admissions and control for the main effect of $POST_t$.

The γ coefficient in (4) will be positive and significant if raising the two numerical limits spurred ISE immigration from previously constrained countries (i.e., high values of $ISE\%_{ct_0}^{Adm}$). Appendix Table 13 shows this to be true, and economies with high values of $ISE\%_{ct_0}^{Adm}$ become the treatment group regardless of actual responses. From the waiting list and 1983-1990 flow data presented in Appendix Table 13, the treated groups are determined to be India, Mainland China, the Philippines, and Taiwan.[12] The reduced-form immigration estimator then takes the form

$$\ln(IMM_{cit}^{RF}) = \ln\left[\sum_{s=1}^{5}(QUOTA_{c,t-s}^{Eff} + QUOTA_{c,t-s-1}^{Eff} + QUOTA_{c,t-s-2}^{Eff})\right], \tag{5}$$

where $QUOTA_{ct}^{Eff}$ is the effective quota for country c in year t. Raising the numerical ceilings did not change the effective quotas for nations unconstrained by the former immigration regime (i.e., low $ISE\%_{ct_0}^{Adm}$), and their effective quotas are held constant at the pre-reform theoretical limit. For constrained countries with high $ISE\%_{ct_0}^{Adm}$ values, the effective quota increases to reflect both the higher country limit of 25,600 and the larger employment preference allocation of 36% (i.e., 120,120/336,000). This quota increase occurs in 1991, and the shift is moved forward to 1990 for Mainland China to account for the CSPA.

5.3 Reduced-Form Results

The robustness of the Full Sample results to excluding India is important. The INS quotas design does not consider shifts in the US ethnic populations of temporary workers (e.g., the H-1B program). The temporary visa program up to the mid-1990s looked quite different from today. Fewer visas were issued, and the most significant occupation and country were medical professionals and the Philippines, respectively. An explosion in Indian temporary workers, mostly for systems analysis and computer programming jobs, began in the 1990s. From 1989-1999, India's share of temporary visas issued rose from 9% to 48% (e.g., Lowell 2000).

Temporary visas can only be renewed once, for a maximum stay of six years, so long-term growth in ethnic research communities requires permanent immigration. Outside of India, the trends for temporary visas are fairly stable for the period studied, and the science and engineering component appears small compared to permanent residency changes. While the jump in India's temporary visa community could affect the final few years of the 1985-1997 period, the results do not depend on its inclusion.

[12] Hong Kong is not included in the treatment group as its immigration status was not affected by the 1990 reform. The main results are robust to instead defining the treatment group at the ethnicity level, although the additional variation inherent in the country-level approach enhances performance in falsification exercises.

References

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

[2] Bartelsman, Eric, and Wayne Gray, "The NBER Manufacturing Productivity Database", NBER Technical Working Paper 205 (1996).

[3] Branstetter, Lee, and Yoshiaki Ogura, "Is Academic Science Driving a Surge in Industrial Innovation? Evidence from Patent Citations", NBER Working Paper 11561 (2005).

[4] Griliches, Zvi, "Productivity, R&D, and the Data Constraint", *The American Economic Review*, 84:1 (1994), 1-23.

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

[6] Harris, John, and Michael Todaro, "Migration, Unemployment, and Development: A Two-Sector Analysis", *The American Economic Review*, 60:1 (1970), 126-142.

[7] Henderson, Rebecca, Adam Jaffe, and Manuel Trajtenberg, "Patent Citations and the Geography of Knowledge Spillovers. A Reassessment: Comment", *The American Economic Review*, 95:1 (2005).

[8] Jaffe, Adam, and Joshua Lerner, *Innovation and its Discontents* (Princeton, NJ: Princeton University Press, 2004).

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

[10] Jasso, Guillermina, Mark Rosenzweig, and James Smith, "The Changing Skill of New Immigrants to the United States: Recent Trends and Their Determinants", NBER Working Paper 6764 (1998).

[11] Johnson, Daniel, "150 Years of American Invention: Methodology and a First Geographic Application", Wellesley College Economics Working Paper 99-01 (1999). Data currently reside at http://faculty1.coloradocollege.edu/~djohnson/uships.html.

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

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

[14] Kerr, William, "Ethnic Scientific Communities and International Technology Diffusion", Working Paper (2007a).

[15] Kerr, William, "Heterogeneous Technology Diffusion and Ricardian Trade Patterns", Working Paper (2007b).

[16] Kerr, William, "The Ethnic Composition of US Inventors", Working Paper (2007c).

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

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

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

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

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

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

[23] Sun, Ling, Lilyan Fulginiti, and Wesley Peterson, "Accounting for Agricultural Decline with Economic Growth in Taiwan", Working Paper (2003).

[24] Thompson, Peter, "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations", *Review of Economics and Statistics*, 88:2 (2006), 383-388.

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

[26] Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data* (Cambridge, MA: MIT Press, 2002).

Figure A1: Transition Path Simulations for Technology Follower

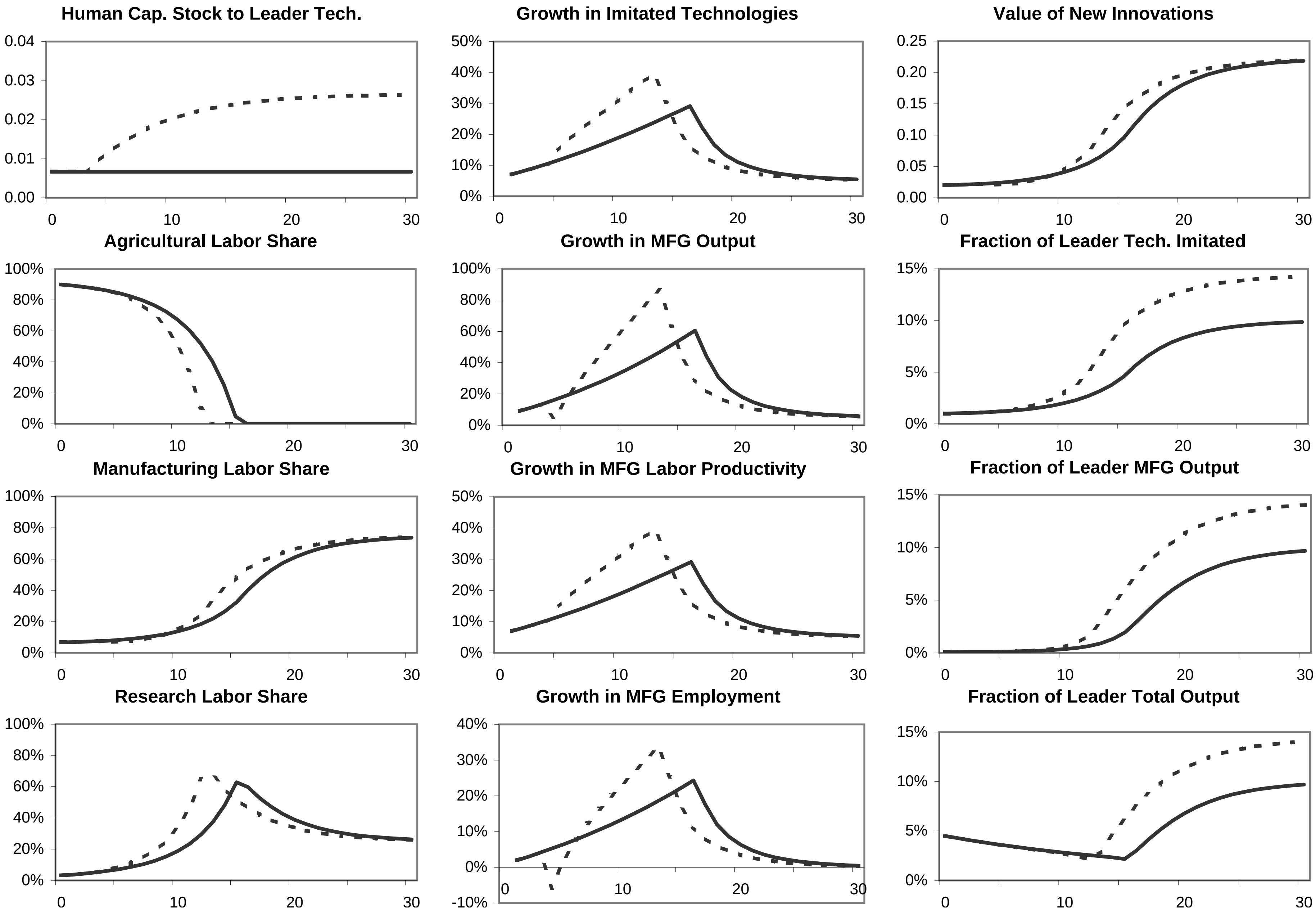

Figure A2: Low-Skilled Immigration

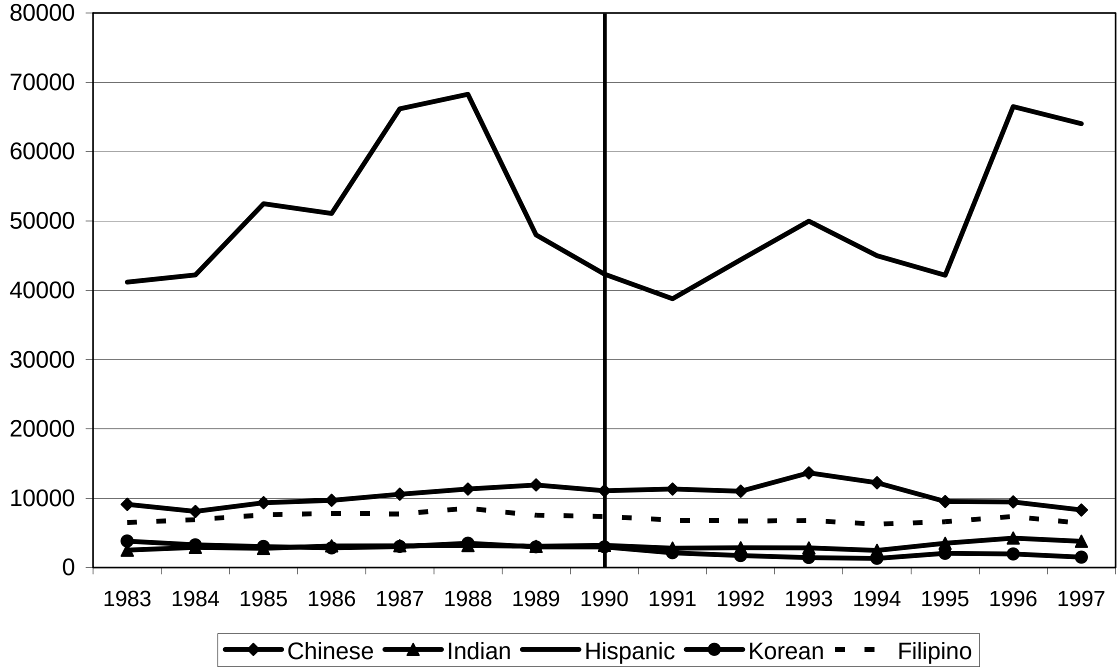

Figure A3: US SE Ph.D. Graduates

Graduates by Country of Origin

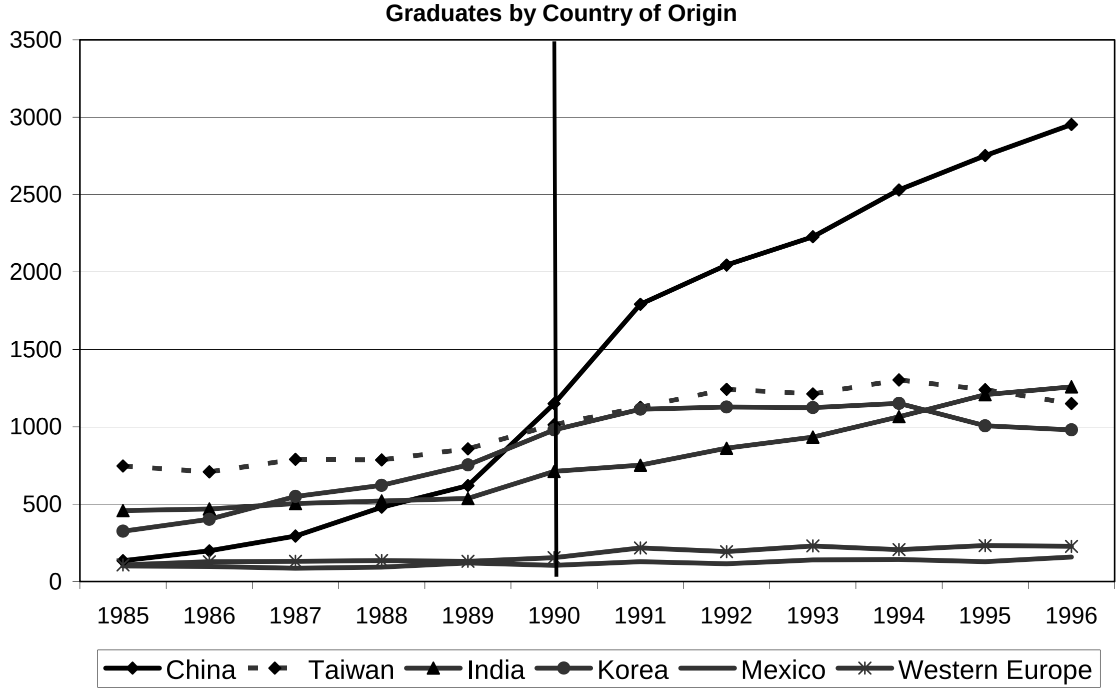

Table A1: Incidence Rate Ratios for Foreign Patent Citations of US Domestic Patents

	Cell Count	All Citations	Citations Separated by Lagged Years from US Domestic Patent Filing to Citing Foreign Patent Filing									
			0-2 Yr.	2-3 Yr.	3-4 Yr.	4-5 Yr.	5-6 Yr.	6-7 Yr.	7-8 Yr.	8-9 Yr.	9-10 Yr.	10-11 Yr.
			A. All Technology Fields at the Two-Digit Level									
Incidence Rate Ratio	93,312	1.496	1.315	1.402	1.485	1.524	1.332	1.399	1.348	1.320	1.332	1.159
for Same Ethnicity		(0.052)	(0.089)	(0.085)	(0.085)	(0.088)	(0.079)	(0.087)	(0.086)	(0.088)	(0.092)	(0.084)
			B. Chemicals at the Three-Digit Level									
Incidence Rate Ratio	426,888	1.268	1.192	1.507	1.368	1.438	1.370	1.046	0.965	1.174	1.043	0.896
for Same Ethnicity		(0.078)	(0.206)	(0.203)	(0.169)	(0.183)	(0.175)	(0.140)	(0.129)	(0.155)	(0.148)	(0.134)
			C. High-Tech at the Three-Digit Level									
Incidence Rate Ratio	720,000	1.300	1.125	1.277	1.225	1.410	1.151	1.323	1.326	1.272	1.206	1.194
for Same Ethnicity		(0.050)	(0.097)	(0.096)	(0.091)	(0.103)	(0.088)	(0.108)	(0.111)	(0.117)	(0.117)	(0.124)
			D. Mechanical at the Three-Digit Level									
Incidence Rate Ratio	943,920	1.234	2.087	1.220	1.503	1.395	1.176	1.092	1.347	1.060	1.363	1.047
for Same Ethnicity		(0.080)	(0.300)	(0.174)	(0.190)	(0.186)	(0.160)	(0.154)	(0.190)	(0.158)	(0.204)	(0.165)
			E. Miscellaneous at the Three-Digit Level									
Incidence Rate Ratio	994,032	1.217	1.702	1.885	1.651	1.505	1.218	1.202	1.501	1.197	1.212	0.959
for Same Ethnicity		(0.081)	(0.237)	(0.237)	(0.198)	(0.184)	(0.152)	(0.167)	(0.207)	(0.166)	(0.177)	(0.147)
			F. Average Lag Structure for Panels B-E									
		1.267	1.468	1.334	1.366	1.415	1.232	1.154	1.213	1.169	1.204	1.045

Notes: Negative binomial regressions contain cells grouping US international patent citations by the ethnicity of citing foreign inventor, the ethnicity of cited US inventor, the technology class of citing foreign inventor, and the technology class of cited US inventor. Regressions are unweighted and include an indicator variable for same ethnicity, an indicator variable for same technology, and vectors of fixed effects for the four dimensions on which cells are constructed. Coefficients on same-ethnicity indicator variable are transformed into the reported incident rate ratios.

Table A2: Thompson (2006) Specifications

	Base Regression	Adding Institutional and Age Controls	Interacting Cited Patent Age
	(1)	(2)	(3)
	Dependent Variable is Same Ethnicity Indicator		
Indicator Variable for Inventor-Added Citation	1.557 (1.70)	1.434 (1.37)	1.758 (1.83)
Interaction of Indicators for Inventor-Added Citations and Cited Patent Over 10 Yrs Old			0.603 (-1.25)
Indicator Variable for Non-Institutional Cited Patent		0.807 (-0.82)	0.777 (-0.97)
Cited Patent Age at Time of Citation		1.085 (2.23)	1.166 (2.66)
Indicator Variable for Cited Patent Being Over 10 Yrs Old			0.884 (-0.29)
Observations	889	889	889

Notes: Table reports conditional logit estimations using Thompson (2006) dataset and technique. Odds ratios are presented with z-scores in parentheses. Sample includes only USPTO patents filed from outside of the US and their citations of patents that were filed inside of the US. The 889 observations are the effective number of observations after the citing patent fixed effects are estimated (from a total sample of 2973).

Table A3: UNIDO Industry Sample

Country	1980 Agr. Share	UNIDO3 Panel	Output (m) Level	Output (m) Growth	Labor Prod. (k) Level	Labor Prod. (k) Growth	Employment (k) Level	Employment (k) Growth	Capital (m) Level	Capital (m) Growth
Single Ethnic Mappings:										
India	70%	85-97	117,950	6%	16	3%	7,354	2%	46,740	4%
Japan	11%	85-97	2,053,048	7%	206	8%	9,998	-1%	415,195	8%
South Korea	37%	85-97	230,942	14%	88	13%	2,626	1%	88,873	14%
Russia	16%	93-97	109,729	12%	10	22%	11,685	-8%		
Soviet Union	16%	85-89	1,087,914	7%	35	8%	31,434	-1%		
Chinese Economies:										
China, Mainland	74%	85-97	327,173	11%	8	9%	38,940	3%		
Hong Kong	1%	85-97	30,520	3%	66	12%	535	-9%	6,628	3%
Macao	6%	85-97	1,209	8%	26	10%	49	-2%	235	1%
Singapore	2%	85-97	37,830	16%	117	12%	309	3%	8,477	8%
Taiwan	8%	85-96	145,055	11%	68	11%	2,141	0%		
European Economies:										
Austria	10%	85-97	73,524	5%	125	5%	595	0%	22,001	5%
Belgium	3%	85-92, 95-97	31,958	5%	131	7%	247	-2%	19,809	7%
Denmark	7%	85-91	38,198	9%	93	11%	411	-1%	8,788	7%
Finland	12%	85-97	52,510	4%	141	8%	386	-4%	18,868	1%
France	8%	85-96	517,276	8%	130	10%	4,006	-2%	107,758	4%
Germany	7%	91-97	870,625	7%	147	7%	5,920	0%		
Germany, East		85-92	233,905	12%	81	12%	2,902	0%		
Germany, West		85-89	734,523	12%	115	12%	6,391	0%	51,571	-6%
Italy	13%	85-94, 96-97	390,266	7%	134	7%	2,897	0%	79,391	6%
Luxembourg	5%	85-97	2,952	3%	137	5%	22	-1%	730	1%
Netherlands	6%	85-97	117,868	6%	178	7%	670	-1%	29,146	6%
Norway	8%	85-97	37,467	4%	149	6%	256	-2%	10,402	-1%
Poland	30%	90-97	54,895	6%	21	7%	2,650	-1%	18,749	1%
Sweden	6%	85-97	93,727	6%	140	7%	678	-1%	23,192	4%
Switzerland	6%	86-96	37,827	7%	142	8%	270	-2%		

Table A3: UNIDO Industry Sample (continued)

	1980	UNIDO3	Output (m)		Labor Prod. (k)		Employment (k)		Capital (m)	
Country	Agr. Share	Panel	Level	Growth	Level	Growth	Level	Growth	Level	Growth
Hispanic Economies:										
Argentina	13%	85-90, 93-96	66,160	11%	73	14%	938	-3%		
Bolivia	53%	85-97	1,474	7%	41	1%	36	6%		
Brazil	37%	90, 92-95	127,807	11%	61	17%	2,105	-5%		
Chile	21%	85-97	20,604	10%	72	5%	278	5%	3,964	9%
Columbia	40%	85-97	20,099	5%	41	3%	487	2%	4,917	-1%
Costa Rica	35%	85-97	3,264	5%	26	1%	127	4%		
Cuba	24%	85-89	10,531	-1%	20	-3%	524	2%	6,097	0%
Ecuador	40%	85-97	4,372	3%	41	2%	107	2%	2,797	1%
Honduras	57%	90-95	989	8%	12	-10%	90	22%		
Mexico	36%	85-97	61,612	4%	60	6%	1,021	-2%	11,111	2%
Panama	29%	85-94, 96-97	1,468	4%	44	3%	33	1%	445	-3%
Peru	40%	85-92, 94-96	13,944	8%	55	9%	255	-1%	2,320	5%
Philippines	52%	85-97	23,238	11%	27	6%	857	5%	5,512	4%
Portugal	26%	85-97	36,365	8%	43	9%	816	-1%		
Spain	18%	85-97	201,951	8%	108	7%	1,858	2%	35,005	7%
Uruguay	17%	85-97	4,648	6%	37	8%	130	-1%		
Venezuela	15%	85-97	24,174	1%	59	2%	417	0%	13,775	1%

Notes: Values are in 1987 US dollars. Levels and growth rates are unweighted averages of yearly country-level aggregates derived from the industry data used in the UNIDO3 panel. Belize, Dominican Republic, El Salvador, Guatemala, Latvia, Lithuania, Nicaragua, Paraguay, and Vietnam are not included due to lack of data. For countries in the sample, insufficient observations or severe quality concerns excluded observations in Bolivia (353 in 1985, 355 and 382 in 1987), Brazil (1985), Costa Rica (371, 385 in 1997), Ecuador (352 in 1994, 354 in 1995, 313 in 1997), Honduras (1981-1989), Hong Kong (369 in 1996), Macao (314) and Venezuela (314 in 1996, 371 in 1995). Series breaks are modeled for Argentina (1990), Austria (1985), China (1989), Denmark (1989), Italy (1994), Mexico (1993), and Portugal (1989) for distinct levels shifts over the 1985-1997 period usually due to changes in variable definitions.

ISIC Rev. 2 Industries: Food products (311), Beverages (313), Tobacco (314), Textiles (321), Wearing apparel, except footwear (322), Leather products (323), Footwear, except rubber or plastic (324), Wood products, except furniture (331), Furniture, except metal (332), Paper and products (341), Printing and publishing (342), Industrial chemicals (351), Other chemicals (352), Petroleum refineries (353), Misc. petroleum and coal products (354), Rubber products (355), Plastic products (356), Pottery, china, earthenware (361), Glass and products (362), Other non-metallic mineral products (369), Iron and steel (371), Non-ferrous metals (372), Fabricated metal products (381), Machinery, except electrical (382), Machinery, electric (383), Transport equipment (384), Professional & scientific equipment (385), and Other manufactured products (390). Industry 390 is excluded.

Table A4: UNIDO Levels Specifications

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Log Foreign Output		
Log US Ethnic	0.241	0.420	0.400
Research Community	(0.126)	(0.228)	(0.147)
Observations	9912	9912	9912
	B. Log Foreign Labor Productivity		
Log US Ethnic	0.215	0.383	0.310
Research Community	(0.088)	(0.180)	(0.124)
Observations	9912	9912	9912
	C. Log Foreign Employment		
Log US Ethnic	0.026	0.037	0.090
Research Community	(0.138)	(0.199)	(0.139)
Observations	9912	9912	9912

Notes: Row titles document the dependent variable studied; column titles document the weighting scheme employed. Panel estimations consider country-industry-year observations taken from the 1985-1997 UNIDO manufacturing database. Log US Ethnic Research Communities are estimated at the ethnicity-industry-year level from the US ethnic patenting dataset. Regressions include country-industry and industry-year fixed effects. Standard errors are conservatively clustered at the ethnicity-industry level.

Table A5: UNIDO Country Controls Specifications

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Base Foreign Output Regression		
Log US Ethnic Research Community	0.241	0.420	0.400
	(0.126)	(0.228)	(0.147)
Observations	9912	9912	9912
	B. Including Foreign Ph.D. Students in US		
Log US Ethnic Research Community	0.225	0.294	0.337
	(0.251)	(0.214)	(0.248)
Log Foreign Ph.D. Students in US	0.016	0.100	0.054
	(0.087)	(0.084)	(0.091)
Observations	8914	8914	8914
	C. Including Foreign Physical-Capital Stocks		
Log US Ethnic Research Community	0.100	0.124	0.190
	(0.176)	(0.275)	(0.199)
Log Foreign Capital Stock	0.432	0.704	0.500
	(0.072)	(0.074)	(0.069)
Observations	5604	5604	5604
	D. Including Country Time Trends		
Log US Ethnic Research Community	0.024	0.078	0.090
	(0.113)	(0.193)	(0.129)
Observations	9912	9912	9912
	E. Including Country-Year Effects		
Log US Ethnic Research Community	0.021	0.243	0.094
	(0.112)	(0.274)	(0.154)
Observations	9912	9912	9912

Notes: See Table A4. Panel A replicates the foreign country-industry output regressions from Table A4. Panels B through E incorporate the country controls indicated by the row titles. All regressions maintain country-industry and industry-year fixed effects and the clustering of standard errors.

Table A6: UNIDO Sample Decompositions

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Base Foreign Output Regression		
Log US Ethnic Research Community	0.241	0.420	0.400
	(0.126)	(0.228)	(0.147)
Observations	9912	9912	9912
	B. Excluding Computers and Drugs		
Log US Ethnic Research Community	0.172	0.087	0.240
	(0.134)	(0.136)	(0.114)
Observations	9067	9067	9067
	C. Excluding Mainland China		
Log US Ethnic Research Community	0.206	0.404	0.382
	(0.146)	(0.269)	(0.172)
Observations	9653	9653	9653
	D. Excluding All Chinese Economies		
Log US Ethnic Research Community	0.379	0.379	0.445
	(0.099)	(0.311)	(0.174)
Observations	8669	8669	8669
	E. Excluding All Advanced Economies		
Log US Ethnic Research Community	-0.056	0.401	0.204
	(0.199)	(0.199)	(0.176)
Observations	6259	6259	6259
	F. Excluding All Hispanic Economies		
Log US Ethnic Research Community	0.252	0.467	0.436
	(0.156)	(0.267)	(0.180)
Observations	5453	5453	5453

Notes: See Table A4. Panel A replicates the foreign country-industry output regressions from Table A4. Panels B through F exclude the observations indicated by the row titles. All regressions maintain country-industry and industry-year fixed effects and the clustering of standard errors.

Table A7: UNIDO Sector Reallocation Specifications

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Log Foreign Output		
Log US Ethnic Research Community	0.237	0.431	0.414
	(0.128)	(0.233)	(0.154)
Log US Ethnic Comm. x 1980 Agriculture Share	0.667	0.340	0.606
	(0.237)	(0.394)	(0.293)
Observations	9912	9912	9912
	B. Log Foreign Labor Productivity		
Log US Ethnic Research Community	0.220	0.360	0.294
	(0.065)	(0.133)	(0.094)
Log US Ethnic Comm. x 1980 Agriculture Share	-0.855	-0.665	-0.654
	(0.102)	(0.172)	(0.130)
Observations	9912	9912	9912
	C. Log Foreign Employment		
Log US Ethnic Research Community	0.017	0.072	0.120
	(0.115)	(0.166)	(0.107)
Log US Ethnic Comm. x 1980 Agriculture Share	1.521	1.005	1.260
	(0.207)	(0.316)	(0.237)
Observations	9912	9912	9912

Notes: Row titles document the dependent variable studied; column titles document the weighting scheme employed. Panel estimations consider country-industry-year observations taken from the 1985-1997 UNIDO manufacturing database. 1980 Agriculture Shares for foreign countries are listed in Table 2. Log US Ethnic Research Communities are estimated at the ethnicity-industry-year level from the US ethnic patenting dataset. Main effects are demeaned prior to interactions. Regressions include country-industry and industry-year fixed effects. Standard errors are conservatively clustered at the ethnicity-industry level.

Table A8: Immigration Quotas Specifications

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Log Foreign Output		
Log US Immigration Quotas Estimator	0.360	0.419	0.368
	(0.156)	(0.281)	(0.198)
Observations	9912	9912	9912
	B. Log Foreign Labor Productivity		
Log US Immigration Quotas Estimator	-0.024	0.039	0.011
	(0.216)	(0.240)	(0.215)
Observations	9912	9912	9912
	C. Log Foreign Employment		
Log US Immigration Quotas Estimator	0.384	0.380	0.357
	(0.081)	(0.085)	(0.074)
Observations	9912	9912	9912

Notes: Row titles document the dependent variable studied; column titles document the weighting scheme employed. Panel estimations consider country-industry-year observations taken from the 1985-1997 UNIDO manufacturing database. Log US Immigration Quotas Estimators are developed from quotas changes due to the 1990 Act. Regressions include country-industry and industry-year fixed effects. Standard errors are conservatively clustered at the ethnicity level.

Table A9: Imm. Quotas Country Controls Specifications

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Base Foreign Output Regression		
Log US Immigration Quotas Estimator	0.360	0.419	0.368
	(0.156)	(0.281)	(0.198)
Observations	9912	9912	9912
	B. Including Foreign Ph.D.s in US		
Log US Immigration Quotas Estimator	0.335	0.362	0.317
	(0.159)	(0.293)	(0.211)
Log Foreign Ph.D. Students in US	0.012	0.131	0.085
	(0.088)	(0.098)	(0.098)
Observations	8914	8914	8914
	C. Excluding Mainland China		
Log US Immigration Quotas Estimator	0.335	0.418	0.358
	(0.191)	(0.342)	(0.243)
Observations	9653	9653	9653
	D. Including Ethnic Time Trend		
Log US Immigration Quotas Estimator	0.502	0.600	0.489
	(0.097)	(0.276)	(0.192)
Observations	9912	9912	9912
	E. Including 1987 Counterfactual		
Log US Immigration Quotas Estimator	0.222	0.265	0.220
	(0.063)	(0.082)	(0.064)
1987 Counterfactual Quotas Estimator	0.206	0.231	0.221
	(0.229)	(0.304)	(0.267)
Observations	9912	9912	9912
	F. Including 1995 Counterfactual		
Log US Immigration Quotas Estimator	0.347	0.514	0.409
	(0.217)	(0.392)	(0.284)
1995 Counterfactual Quotas Estimator	0.036	-0.260	-0.113
	(0.180)	(0.316)	(0.240)
Observations	9912	9912	9912

Notes: See Table A8. Panel A replicates the foreign country-industry output regressions from TableA8. Panels B through F incorporate the country controls indicated by the row titles. All regressions maintain country-industry and industry-year fixed effects and the clustering of standard errors.

Table A10: UNIDO Capital-Labor - Levels

	No Weights	Patent Weights	Output Weights
	(1)	(2)	(3)
	A. Base Foreign Productivity Regression		
Log US Ethnic Research Community	0.215	0.383	0.310
	(0.088)	(0.180)	(0.124)
Observations	9912	9912	9912
	B. Restricted Capital Sample		
Log US Ethnic Research Community	0.132	0.445	0.295
	(0.131)	(0.251)	(0.172)
Observations	5604	5604	5604
	C. Including Capital-Labor Ratio		
Log US Ethnic Research Community	0.121	0.332	0.238
	(0.121)	(0.241)	(0.159)
Log Foreign Capital-Labor Ratio	0.212	0.244	0.216
	(0.061)	(0.051)	(0.055)
Observations	5604	5604	5604

Notes: See Table A4. Panel A replicates the foreign country-industry labor productivity regressions (levels) from Table A4. Panel B re-estimates the labor productivity regressions for observations with capital data. Panel C incorporates the foreign country-industry capital-labor ratios. Regressions also maintain industry-year and country-industry fixed effects.

Table A10: UNIDO Capital-Labor - First-Differences

	No Weights	Patent Weights	Output Weights
	(4)	(5)	(6)
	A. Base Foreign Productivity Regression		
Δ Log US Ethnic Research Community	0.087	0.214	0.217
	(0.049)	(0.114)	(0.072)
Observations	8736	8736	8736
	B. Restricted Capital Sample		
Δ Log US Ethnic Research Community	0.044	0.194	0.166
	(0.062)	(0.144)	(0.087)
Observations	4866	4866	4866
	C. Including Capital-Labor Ratio		
Δ Log US Ethnic Research Community	0.049	0.191	0.164
	(0.061)	(0.144)	(0.085)
Δ Log Foreign Capital-Labor Ratio	0.111	0.033	0.083
	(0.028)	(0.042)	(0.034)
Observations	4866	4866	4866

Notes: See Table 3. Panel A replicates the foreign country-industry labor productivity regressions (first-differences) from Table 3. Panel B re-estimates the labor productivity regressions for observations with capital data. Panel C incorporates the foreign country-industry capital-labor ratios. Regressions also maintain industry-year fixed effects.

Table A11: UNIDO Sector Reallocation Regressions - Levels

				Patent Weights				
	No Weights	Patent Weights	Output Weights	Excluding Computers and Drugs	Excluding Mainland China	Excluding All Chinese Economies	Excluding All Advanced Economies	Excluding All Hispanic Economies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Log Foreign Output				
Log US Ethnic Research Community	0.237	0.431	0.414	0.105	0.431	0.034	0.512	0.463
	(0.128)	(0.233)	(0.154)	(0.153)	(0.263)	(0.367)	(0.249)	(0.261)
Log US Ethnic Comm. x 1980 Agriculture Share	0.667	0.340	0.606	0.653	0.343	0.810	0.307	0.076
	(0.237)	(0.394)	(0.293)	(0.262)	(0.414)	(0.569)	(0.440)	(0.402)
Observations	9912	9912	9912	9067	9653	8669	6259	5453
				B. Log Foreign Labor Productivity				
Log US Ethnic Research Community	0.220	0.360	0.294	0.261	0.294	-0.153	0.772	0.399
	(0.065)	(0.133)	(0.094)	(0.164)	(0.122)	(0.287)	(0.150)	(0.090)
Log US Ethnic Comm. x 1980 Agriculture Share	-0.855	-0.665	-0.654	-0.707	-0.876	-0.362	-0.512	-0.510
	(0.102)	(0.172)	(0.130)	(0.202)	(0.188)	(0.280)	(0.211)	(0.246)
Observations	9912	9912	9912	9067	9653	8669	6259	5453
				C. Log Foreign Employment				
Log US Ethnic Research Community	0.017	0.072	0.120	-0.155	0.138	0.187	-0.260	0.063
	(0.115)	(0.166)	(0.107)	(0.094)	(0.188)	(0.186)	(0.162)	(0.185)
Log US Ethnic Comm. x 1980 Agriculture Share	1.521	1.005	1.260	1.360	1.220	1.172	0.818	0.586
	(0.207)	(0.316)	(0.237)	(0.257)	(0.315)	(0.391)	(0.321)	(0.268)
Observations	9912	9912	9912	9067	9653	8669	6259	5453

Notes: Row titles document the dependent variable studied; column titles document the weighting scheme employed. Panel estimations consider country-industry-year observations taken from the 1985-1997 UNIDO manufacturing database. 1980 Agriculture Shares for foreign countries are listed in Table A3. Log US Ethnic Research Communities are estimated at the ethnicity-industry-year level from the US ethnic patenting dataset. Main effects are demeaned prior to interaction. Regressions include industry-year and country-industry fixed effects. Standard errors are conservatively clustered at the ethnicity-industry level.

Table A12: UNIDO Sector Reallocation Regressions - First-Differences

	No Weights	Patent Weights	Output Weights	Patent Weights: Excluding Computers and Drugs	Patent Weights: Excluding Mainland China	Patent Weights: Excluding All Chinese Economies	Patent Weights: Excluding All Advanced Economies	Patent Weights: Excluding All Hispanic Economies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Δ Log Foreign Output								
Δ Log US Ethnic Research Community	0.043	0.315	0.252	0.083	0.315	0.019	0.425	0.295
	(0.062)	(0.153)	(0.086)	(0.093)	(0.176)	(0.129)	(0.126)	(0.202)
Δ Log US Ethnic Comm. x 1980 Agriculture Share	0.765	0.442	0.647	0.754	0.322	0.793	0.349	0.283
	(0.185)	(0.353)	(0.242)	(0.189)	(0.351)	(0.379)	(0.367)	(0.376)
Observations	8736	8736	8736	7991	8518	7616	5549	4821
B. Δ Log Foreign Labor Productivity								
Δ Log US Ethnic Research Community	0.105	0.225	0.228	0.109	0.173	-0.083	0.498	0.219
	(0.047)	(0.106)	(0.068)	(0.105)	(0.096)	(0.117)	(0.099)	(0.096)
Δ Log US Ethnic Comm. x 1980 Agriculture Share	-0.284	-0.191	-0.216	-0.167	-0.572	-0.169	-0.237	-0.053
	(0.097)	(0.162)	(0.120)	(0.199)	(0.158)	(0.188)	(0.164)	(0.211)
Observations	8736	8736	8736	7991	8518	7616	5549	4821
C. Δ Log Foreign Employment								
Δ Log US Ethnic Research Community	-0.062	0.091	0.024	-0.026	0.142	0.102	-0.073	0.076
	(0.037)	(0.084)	(0.047)	(0.048)	(0.103)	(0.094)	(0.061)	(0.117)
Δ Log US Ethnic Comm. x 1980 Agriculture Share	1.049	0.633	0.863	0.922	0.894	0.962	0.586	0.336
	(0.146)	(0.266)	(0.198)	(0.187)	(0.259)	(0.302)	(0.282)	(0.234)
Observations	8736	8736	8736	7991	8518	7616	5549	4821

Notes: Row titles document the dependent variable studied; column titles document the weighting scheme employed. Panel estimations consider country-industry-year observations taken from the 1985-1997 UNIDO manufacturing database. 1980 Agriculture Shares for foreign countries are listed in Table A3. Log US Ethnic Research Communities are estimated at the ethnicity-industry-year level from the US ethnic patenting dataset. Main effects are demeaned prior to interactions. Regressions include industry-year fixed effects. Standard errors are conservatively clustered at the ethnicity-industry level.

Table A13: Immigration Quotas Reduced-Form Preliminaries

	Regressions of Immigration Response to 1990 Act (Thousands) By Occupation				Regressions of Log US Ethnic Patents on Log US Immigration Quotas		
	Scientists	Business	Total		No Weights	Patent Weights	Output Weights
1983-1990 Quota Share x Post 1990	4.669	4.842	3.279	Log US Immigration Quotas Estimator	0.217	0.256	0.213
	(0.183)	(0.127)	(0.125)		(0.109)	(0.118)	(0.103)
Observations	2310	2310	2310	Observations	9912	9912	9912

	1983-1990 Percent of Theoretical Employment Quota for Country				Employment Visa Waiting List January 1992		
	Scientists	Business	Total		High-Skill	Skilled	Low-Skill
Hong Kong	20.5%	15.6%	102.6%	The Philippines	6795	9550	5995
India	18.5%	5.7%	83.3%	Mainland China	3266	1942	2976
Taiwan	18.2%	10.8%	102.0%	India	3132	1156	1131
United Kingdom	11.7%	13.9%	103.7%	Taiwan	2065	2411	1613
Iran	8.4%	4.5%	54.1%	Nigeria	1854	166	298
Mainland China	6.5%	5.3%	57.1%	Great Britain	1841	2521	714
The Philippines	4.6%	8.4%	96.4%	Canada	1587	2107	191
Canada	3.8%	9.5%	67.7%	Hong Kong	811	1350	885
South Korea	2.2%	5.0%	69.0%	Iran	804	1536	927
Pakistan	1.8%	1.4%	13.0%	Japan	787	1634	800
Israel	1.7%	1.6%	24.5%	South Korea	539	1656	5466
World Average	0.8%	0.8%	8.8%	Total	50,003	32,452	87,806

Notes: Immigration response estimations test immigration responses of quotas-constrained countries to the 1990 Act by occupation. US permanent residency admissions are regressed on each country's 1983-1990 occupation admissions divided by the theoretical country-employment limit interacted with an indicator variable for post the 1990 Act's effective date. Regressions include country and year effects, cluster standard errors at the country level, and exclude Hong Kong due to its special US immigration treatment. The 1983-1990 quota shares are documented in the bottom-left panel for major countries; the bottom-right panel documents INS waiting list records close to the October 1991 effective date of the 1990 Act.

Patent regressions test US ethnic invention responses to 1990 Act. Ethnic patenting is regressed on the constructed immigration quotas estimators. Regressions include country-year and industry-year effects; standard errors are clustered at the ethnicity