Women's empowerment and the return migration of skilled workers

Alireza Javadian Sabet[a,b], **Junghyun Lim**[c], **Yixiao Li**[a], **Lisa K. Simon**[d], **and Morgan R. Frank**[a,b,e,1]

This manuscript was compiled on July 5, 2026

The rapidly growing knowledge economy has intensified the global competition for talent, particularly among emerging economies that invest heavily in attracting skilled workers back from abroad. Women constitute a growing share of both the skilled workforce and migrants seeking education and training. Despite this trend, we know relatively little about gender differences in the return migration patterns of highly skilled migrants. To what extent do gender gaps exist in the international mobility of highly skilled workers, and what may predict differences in migration patterns by gender? We focus on women's empowerment in countries of origin, capturing women's political rights, representation, and access to social and economic opportunities. Limited rights and opportunities for women may be associated with the persistent exit of female talent and wider gender gaps in return migration by disproportionately increasing the costs of return for women. We analyze millions of LinkedIn profiles to trace the education, career trajectories, and migration patterns of immigrants to the U.S. Compared to male compatriots, women from countries with lower women's empowerment are less likely to return to their origin countries. This pattern holds across multiple indicators including female representation in government, labor-force participation, and educational attainment. Migrants who remain abroad tend to publish more peer-reviewed articles than returnees and otherwise similar non-migrants in their countries of origin. Taken together, these findings suggest that women's rights in countries of origin are linked to persistent "brain drain" and highlight economic incentives for emerging economies to improve women's rights.

Return migration | Brain drain | Gender equality | Skilled migration | Scientific productivity

Significance Statement

As women comprise a growing share of international students and workers, global competition for talent increasingly intersects with social institutions and gendered opportunity structures. Using millions of career histories, we show that women's empowerment—laws, opportunities, and social norms that expand women's agency—is associated with skilled migrants' likelihood of returning home after working or training abroad. Women from countries with lower empowerment are less likely to return, while migrants from these countries are also more likely to pursue U.S. doctorates. Women from countries with higher empowerment are more likely to return, and gender gaps in scientific productivity are smaller in these settings. Taken together, these findings suggest that domestic social structures are linked to "brain drain" and highlight women's empowerment as a potentially important lever for countries seeking to encourage a more gender-balanced circulation of talent.

International migration shapes global knowledge production, technology transfer, and economic development. Transnational individuals often raise their own productivity (1–6) and that of others in their host and home countries by acquiring and diffusing new skills (7–11). Skilled immigrants constitute a major share of the U.S. innovation workforce (12–15) with foreign-born workers accounting for roughly 25% of the science and engineering (S&E) workforce, close to 50% of S&E workers with PhDs, and more than half of founders of billion-dollar startups (16, 17).

As the knowledge economy rapidly expands, competition for highly skilled workers also accelerates (18). Many developing countries focus not only on sending individuals abroad for training and education but also on repatriating them (19). In this global competition, women comprise a growing share of the skilled workforce. UNESCO estimates that women account for about 53% of bachelor's and master's graduates and 43% of PhD graduates worldwide (20). Growing female representation also emerges in research publications (21), in STEM fields, where women have historically been underrepresented (22), and among internationally mobile populations seeking education and training abroad (23).

This trend suggests the importance of repatriating highly skilled women from abroad in the global competition for talent, particularly for emerging economies. Losing a cohort of highly trained women drains the skilled talent domestically and exacerbates existing gender gaps in research and innovation (21, 24). The departure of female scientists and entrepreneurs may limit the gender diversity vital to research innovation and institutional reform (25–27). Highly skilled women who return from abroad often bring professional experiences and expanded international networks that contribute to local job creation, technological diffusion, and mentorship initiatives, potentially fostering a longer term growth in the female workforce (28–31).

Despite the growing importance of repatriating female talent, we know relatively little about gender differences in the return migration patterns of highly skilled migrants. While an increasing number of studies recognize the importance of gender as a predictor of migration flows (32), fewer studies focus on highly skilled

Author affiliations: [a]Department of Informatics and Networked Systems, University of Pittsburgh, Pittsburgh, PA 15215, USA; [b]Digital Economy Lab, Stanford University, Stanford, CA 94305, USA; [c]Department of Political Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; [d]Revelio Labs, New York, NY 10010, USA; [e]AI Economy Institute, Microsoft, Redmond, WA 98052, USA

A.J.S. prepared the LinkedIn data, performed calculations and produced figures. M.R.F. secured funding for this project. Y.L. prepared the publication data from SciSciNet. L.K.M. secured resume data from Revelio Labs. J.L. and M.R.F conceptualized the study. A.J.S., J.L., and M.R.F. designed the research. All authors drafted the manuscript.

The authors declare no competing interests.

[1]To whom correspondence should be addressed. E-mail: mrfrank@pitt.edu

DRAFT

women and their return migration (33). To what extent do gender gaps exist in the international mobility of highly skilled workers, and what may predict differences in return migration patterns by gender?

In this paper, we examine gender gaps in return migration of highly skilled workers, with a focus on women's empowerment in countries of origin, capturing women's rights, representation, and access to social and economic opportunities. Classical migration theories with a cost-benefit framework suggest that individuals migrate when the expected benefits outweigh the economic and social costs. By the same logic, individuals are more likely to return when they expect greater benefits from returning to their countries of origin than from remaining abroad. However, the costs of return may vary by gender even among individuals from the same countries of origin with similar economic opportunities. The costs or gains of return may vary by the extent of women's rights in countries of origin. Limited rights for women may be associated with the persistent loss of female talent and wider gender gaps in return migration by disproportionately increasing the costs of return for women.

To examine this, we use millions of LinkedIn profiles to quantify a gender gap in the return migration of highly skilled migrants in the U.S. from 211 countries and territories of origin, and examine how social barriers experienced by women in their countries of origin correspond to return rates. LinkedIn worker profiles provide large-scale, longitudinal records of education and work, together with locations and gender. These profiles describe individuals from their initial migration to the U.S. through to later migration decisions, including returning to their country of origin, remaining in the U.S., or moving elsewhere, and thus enable the study of mobility and potential gender gap in international "brain drain." These data are especially valuable as the gender dimensions of emigration and return migration rarely documented in existing migration datasets (34, 35).

Combining these data, we find that, overall, highly skilled women are less likely to return to their home countries compared to their male compatriots, and this gap is particularly pronounced for countries of origin with lower levels of women's empowerment, leaving a larger share of female talent abroad. Moreover, those who do not return tend to have higher scientific productivity compared to their counterparts who return, regardless of gender.

This study makes three contributions. First, we provide large-scale evidence on gender gaps in return migration among highly skilled migrants, a population that is difficult to observe in existing migration data. Second, we show that these gender gaps are systematically associated with women's empowerment in countries of origin, suggesting that domestic institutions may be an important correlate of who returns after experience abroad. Third, by linking migration trajectories to publication outcomes, we show that non-return is associated with the location of scientific productivity, highlighting a potential channel through which origin countries may lose highly productive talent. Taken together, these findings suggest that lower women's empowerment in countries of origin is associated with persistent "brain drain" and that greater women's empowerment is linked to higher repatriation rates for highly skilled women abroad.

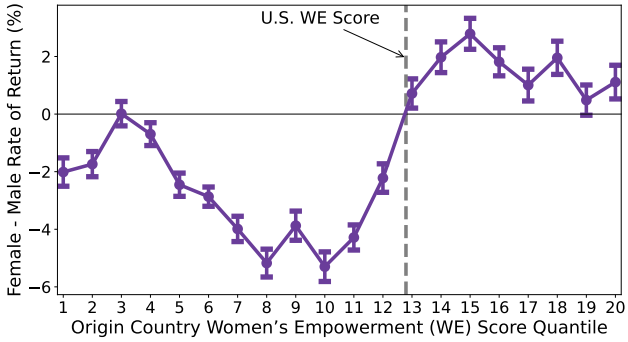


Fig. 1. Female migrants from countries with women's empowerment below the U.S. are less likely than male compatriots to return. Error bars denote 95% confidence intervals for the female–male difference in return rates. The dashed line marks the average U.S. empowerment score from 1990 to 2022.

Results

Women's Empowerment and Female Return Migration. Do women's rights in countries of origin predict highly skilled female emigrants' decisions to return after experiences in the U.S.? To measure women's rights, we use the country-year Women's Empowerment ($WE_{c,y}$) index from the Varieties of Democracy (V-Dem) project*. This index captures women's civil liberties, civic participation, and political representation in 175 countries across the world between 1990 and 2022. In robustness checks, we replace this index with related V-Dem measures comprising the WE and with gender-equality indicators from the United Nations Development Programme (see SI, Sections S1.1.2–S1.1.3). We track U.S. immigrants' education and work experience using 2,537,090 individual LinkedIn profiles provided by Revelio Labs. Each profile includes demographic characteristics (i.e., inferred gender), educational experiences (i.e., degree type, field of study, year of degree completion), and a career history (i.e., job title, employer, location, start and end dates, industry). As we focus on highly skilled immigrants, we focus on individuals who holds at least a Bachelor's or its equivalent. Because the country of origin is not directly observed, we infer it as the country of each individual's first recorded education or employment spell, assuming that this earliest observed location reflects the origin country, and then track subsequent immigration to the U.S. This data also allows us to observe immigrants' later career decisions to remain in the U.S., emigrate elsewhere, or to return to their origin country. Using this data, we can measure the difference in return rates for females and males from the same origin country.

Compared to male compatriots, highly skilled women from countries with lower levels of women's empowerment (WE) than the U.S. are significantly less likely to return to their home countries (see Fig. 1). For emigrants from countries with $WE > WE^{USA}$, women are instead more likely to return to their countries of origin. The same pattern emerges when we adjust for other individual- and country-level covariates, including individuals' education levels or economic performance of origin countries. We estimate the logistic model separately for individuals who first came as students and as those who first migrated as workers to account for potential differences in their intentions to return or stay

*https://www.v-dem.net/

DRAFT

Table 1. Rates of return migration by cohort's women's empowerment. Logistic regressions estimate the probability of returning to the origin country for migrants who first arrived as students or as workers. Each specification includes indicators for gender, women's empowerment (z-score), their interaction, residualized log GDP per capita, years spent in the United States, and prior U.S. job or education experience, with year and origin–region fixed effects. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions. Standard errors are clustered by country of origin.

Dependent Variable: Returned		
	Students	Workers
Female	-0.073**	-0.074*
	(0.027)	(0.041)
Women's Empowerment (z-score)	0.159*	0.286**
	(0.067)	(0.098)
Female*Women's Empowerment (z-score)	0.052*	0.129**
	(0.026)	(0.044)
GDP per Capita (log) (Residualized)	0.329***	0.283***
	(0.066)	(0.086)
Years in the U.S.	-0.233***	-0.319***
	(0.012)	(0.016)
U.S. Job Experience	-0.640***	
	(0.055)	
U.S. Education Experience		-0.613***
		(0.074)
Year FE	Yes	Yes
Region FE	Yes	Yes
Pseudo R^2	0.230	0.258
Adjusted Pseudo R^2	0.229	0.258
Number of Observations	1,051,090	1,486,000
Significance levels: $^*p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$		

from the initial stage of migration (see Table 1). At the sample mean of the standardized women's empowerment index ($WE = 0$), women are less likely to return than men: women's odds of return are about 7% lower than men's in both the student and worker cohorts[†]. Higher levels of women's empowerment in countries of origin is associated with greater return rates for both genders, raising the likelihood of return for everyone but more markedly for women. A one–standard-deviation increase in WE corresponds to about a 17% increase in men's odds of return and a 23% increase for women among those who first arrived as students, and to about 33% and 51% higher odds of return for men and women, respectively, among those who first arrived as workers. We demonstrate this pattern using the predicted probabilities and average marginal effects derived from these logistic regressions (see Fig. 2): return rates rise as the women's empowerment score increases for each cohort and gender. Yet the trajectories for women are steeper. Male migrants are more likely to return to their home countries than their female counterparts when WE_{origin} is low. Yet, the gender gap in return migration rates closes around the level observed in the United States, and at the top of the distribution, female migrants show higher return rates, especially among those who initially migrated for work.

Economic conditions in the origin country are another important factor to account for. Economic conditions likely influence the return migration patterns of highly skilled workers regardless of gender, as they capture potential economic gains and opportunities from return migration (36, 37). In general, however, economic conditions in the origin countries tend to be highly correlated with levels of women's empowerment, creating concerns regarding multicollinearity. To account for multicollinearity and separate the role of prosperity from that of women's rights, we orthogonalize GDP per capita with respect to empowerment and use the residual as a model covariate, following (38). That is, we remove the portion of log GDP per capita that can be linearly predicted from women's empowerment, and include only the remaining component in the regression. Residualization (a.k.a. "partialling out") is a standard way to reduce collinearity and estimate a regressor's association after removing the part explained by a closely related construct; the coefficient then reflects the association with departures from the value predicted by that construct. The residual term captures whether a country is richer or poorer than would be expected given its women's empowerment score, while the coefficient of women's empowerment reflects differences in gender–inclusive institutions at a given level of GDP. Using these specifications, the interaction between gender and women's empowerment score remains significantly positive, suggesting that women are less likely to return to their home countries when their country of origin has low levels of women's empowerment.

Beyond these country-level factors, individual migration histories are also associated with return. Longer stays in the U.S. correspond with lower return rates, consistent with gradual anchoring in the destination country and potential employability issues in origin countries (39) (see Table 1). Prior U.S. job experience among those who arrived as students, and prior U.S. education among those who arrived as workers, both correspond to markedly lower probabilities of going home. These patterns suggest that return migration is also related to migrants' accumulated experience in the U.S. labor market and education system, although these variables should be interpreted as descriptive controls rather than causal estimates.

This pattern remains consistent across a series of robustness checks. First, using alternative measures of women's empowerment drawn from UNDP and from V-Dem subindices yields a highly consistent pattern for both students and workers (see SI, Tables S9 and S10). Across these alternative measures, women remain less likely than men to return overall, yet they are more likely to return when women's rights in their origin country are less restricted. Second, our conclusions do not hinge on *when* empowerment is measured.

Using the empowerment level in the year of first U.S. entry (immigration year) or in the individual's last observed year (i.e., year to decide stay or return) (see SI, Section S1.2 for construction of these variables) produces very similar estimates (see SI, Table S11), with slightly larger point estimates when empowerment is measured at the time the decision to stay or return is made. We describe the construction of the first-entry year and last-observed year measures in SI, Section S1.2. Third, the results are robust to reweighting the sample to better represent the origin-country composition of migrants using DS-visa coverage benchmarks, administrative records of U.S. visa issuances by origin country (see SI, Table S8). LinkedIn profiles may not be representative

[†] For reference, the average U.S. empowerment score over 1990–2022 is higher than the sample mean, with a raw value of 0.898 compared to the sample mean of 0.791, or about 0.73 standard deviations above the sample mean, placing the U.S. average in the 13th of 20 empowerment bins in Fig. 1.

PREPRINT

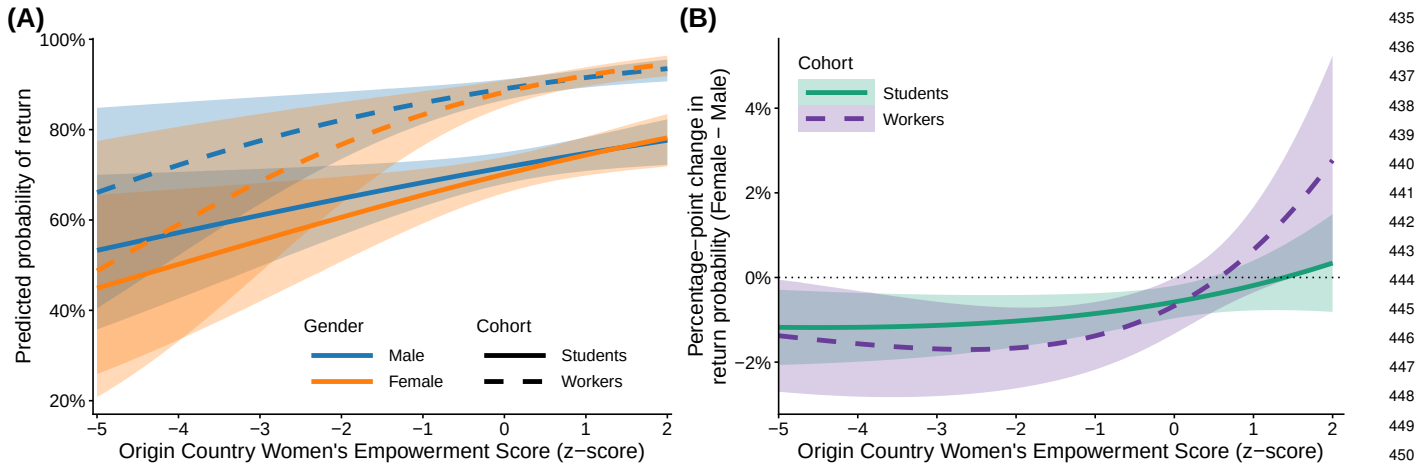


Fig. 2. Return migration rates by gender, cohort, and women's empowerment. Predicted probabilities of return migration are plotted across the Women's Empowerment scores for men and women who first arrived as students (solid) or workers (dashed) in the U.S. **(A)** shows the level of return probability; **(B)** shows the percentage-point difference in return probability for women relative to men (i.e., average marginal effects) with positive values indicating higher return probabilities for women than for men. Estimates are based on logistic regressions in Table 1, which control for Year, Region, Years in the U.S., GDP, and U.S. Job/Education Experience; shaded ribbons indicate 95% confidence intervals.

of all highly skilled migrants in the United States, as migrants from some countries may be over- or underrepresented in the data. To address this concern, we reweight the sample using administrative data.

Applying country-level post-stratification weights, the results remain consistent: women are less likely than men to return on average, yet they are more responsive to the level of women's empowerment in their countries of origin, compared to their male counterparts. This suggests that the central relationship between women's empowerment and gender gap in return migration rates is not driven by over- or underrepresentation of particular origin countries in the sample.

Women's Empowerment and Engagement with U.S. PhD Programs. Next, we examine migrants' educational attainment by gender, focusing on enrollment in U.S. PhD programs. While our sample already consists of highly skilled individuals who have at least a bachelor's degree or its equivalent, doctoral degrees represent advanced research training and typically require an original intellectual contribution, often through a dissertation; PhD-trained researchers also enter research careers in academia, industry R&D, and entrepreneurship (40–42). At the same time, U.S.-trained foreign doctorate recipients often remain in the United States after graduation: recent estimates show that about 70% remain five years after graduation and 62% remain ten years after graduation, while about 76% report plans to stay at the time they receive their degree (43, 44). If female migrants from countries with lower levels of empowerment tend to pursue doctoral degrees in the U.S. more often, countries that fail to repatriate them may disproportionately lose highly trained and innovative workers.

In general, female migrants are less likely than their male counterparts to enroll in a U.S. PhD program among those who first arrived as students; among those who first arrived as workers, the average gender difference is small and not statistically signifciant(Table 2).

Table 2. U.S. PhD enrollment and women's empowerment by cohort. Logistic regressions estimate the probability of enrolling in a U.S. PhD program for migrants who first arrived as students or as workers. Each specification includes indicators for gender, women's empowerment (z-score), their interaction, residualized log **GDP per capita, years in the United States, and prior U.S. job or education experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.**

Dependent Variable: Enrolling in a U.S. Ph.D. Program		
	Students	**Workers**
Female	-0.237***	-0.039
	(0.058)	(0.045)
Women's Empowerment (z-score)	-0.360***	-0.392***
	(0.093)	(0.091)
Female×Women's Empowerment (z-score)	0.012	0.079*
	(0.037)	(0.036)
GDP per Capita (log) (Residualized)	-0.250***	-0.100
	(0.066)	(0.068)
Years in the U.S.	0.090***	0.255***
	(0.007)	(0.015)
U.S. Job Experience	0.304***	
	(0.065)	
U.S. Education Experience		0.937***
		(0.072)
Year FE	Yes	Yes
Region FE	Yes	Yes
Pseudo R^2	0.112	0.121
Adjusted Pseudo R^2	0.112	0.120
Number of Observations	1,051,090	1,486,000
Significance levels: $^*p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$		

Regardless of gender, immigrants from countries with higher women's empowerment were overall less likely to pursue a U.S. PhD (see Fig. 3A). Among migrants who first entered as students, predicted PhD enrollment is highest for those from countries with lower levels of women's empowerment: predicted enrollment is roughly 25% for men and 20% for women at the low end of the empowerment distribution. This gender difference reflects the lower average

DRAFT

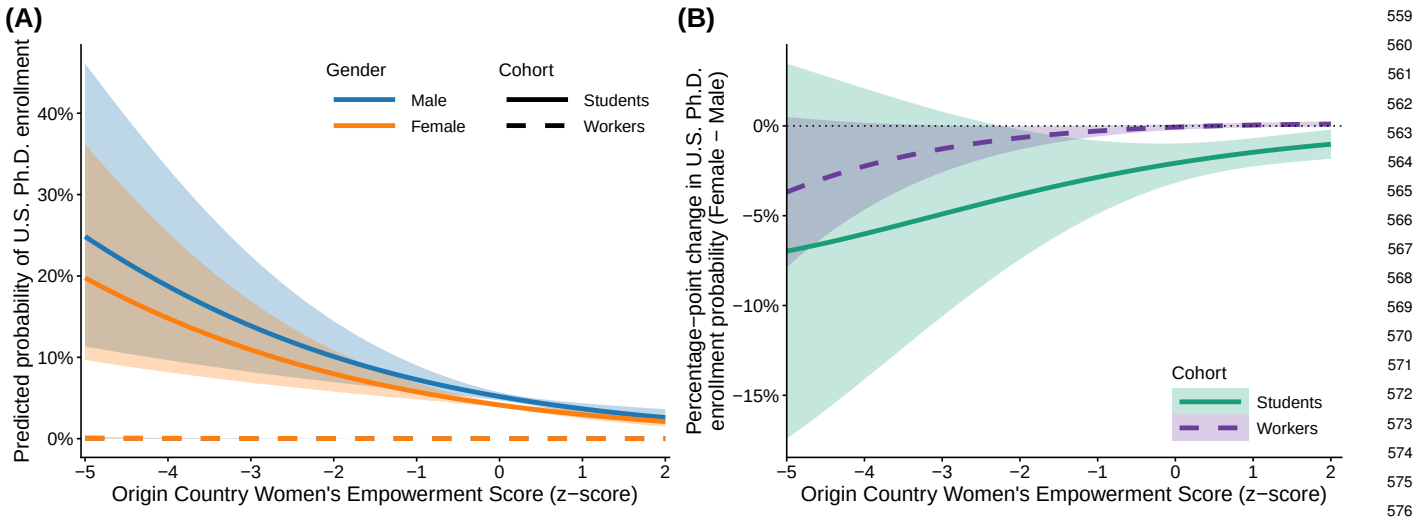


Fig. 3. U.S. PhD enrollment by gender, cohort, and women's empowerment. (A) The level of PhD enrollment probability. **(B)** The percentage-point difference in enrollment probability for women relative to men (average marginal effects) over the same range of empowerment; positive values indicate higher PhD enrollment probabilities for women than for men. Estimates are based on logistic regressions in Table 2, which control for Year, Region, Years in the U.S., GDP, and U.S. Job/Education Experience; shaded regions indicate 95% confidence intervals.

PhD enrollment of women in the student cohort, although the gender-by-empowerment interaction is not statistically significant. For those who first entered the U.S. as workers, the predicted PhD enrollment rate is below 1% for both men and women, reflecting that workers rarely switch into a U.S. PhD. For migrants from countries with higher levels of women's empowerment, predicted PhD enrollment rates fall below 5% for all groups. The decline is steeper for men in the worker cohort, so the gap between male and female PhD enrollment narrows at higher empowerment levels and approaches zero, with a slight reversal at the upper end of the observed range (see Fig. 3B).

We also account for other covariates that may influence migrants' pursuit of a PhD. Among countries with similar levels of empowerment, migrants from relatively richer origins are less likely to enroll in a U.S. PhD, especially among those who first arrive as students. This pattern may reflect broader educational or career opportunities available to migrants from richer origins, reducing reliance on U.S. doctoral programs as a pathway for advancement (45, 46). Longer stays in the United States are positively associated with PhD enrollment, and prior exposure to U.S. institutions matters strongly: for those who arrive as students, U.S. job experience raises the odds of pursuing a PhD, and for those who arrive as workers, prior U.S. education is an especially powerful predictor of subsequent doctoral study. These relationships are stable across alternative model specifications (see SI, Table S13). For example, we obtain similar results when we restrict the student cohort to migrants whose first U.S. degree is not at the PhD level (see SI, Table S14), ruling out the possibility that our conclusions are driven by those who immigrate specifically for a PhD. They are also robust to replacing PhD enrollment with completed U.S. PhD attainment as the outcome (see SI, Table S15), suggesting that the observed patterns reflect not only entry into doctoral training but also degree completion.

Taken together, these results suggest that U.S. PhD enrollment is more common among migrants from lower-empowerment origins, especially among those who first enter the U.S. as students. However, this pattern is not primarily driven by women having higher PhD enrollment than men: in the student cohort, women enroll at lower average rates than men, and the gender-by-empowerment interaction is not statistically significant. Thus, doctoral training appears to be an additional pathway through which highly skilled migrants from lower-empowerment countries may remain attached to the U.S. education and labor-market system.

Women's Empowerment and Immigrants' Scientific Productivity. Compared to male compatriots, female migrants from countries with low women's empowerment are less likely to return. But to what extent permanent exit of female talent contribute to "brain drain" in those countries? We assess how disproportionate exit of female migrants may translate into academic productivity by comparing the scientific output of those who stay abroad with that of those who return. While academic research is not the only indicator of the potential contribution of migrants, it is widely used as an indicator of innovation, and a proxy for productivity and economic growth in knowledge-based economies (19, 23, 47–52). We compare the publication records of three groups stratified by gender: (1) those who never left their home countries (Non-immigrants), (2) those who returned after experiences in the U.S. (Returnees), and (3) those who stay abroad (Stayers).

Across all origin countries, international experience is strongly associated with higher cumulative output. At the mean level of women's empowerment, and accounting for region, field of study, and GDP, immigrants publish about 98% more papers than comparable non-immigrants (see Table 3), while returnees publish about 77% more than comparable non-immigrants in their countries of origin. Among migrants, those who stay in the United States are more productive than those who return, regardless of gender: male stayers publish about 15% more papers than male returnees, and female stayers publish about 14% more than female returnees at the mean level of women's empowerment. The gap of research output between female stayers and

DRAFT

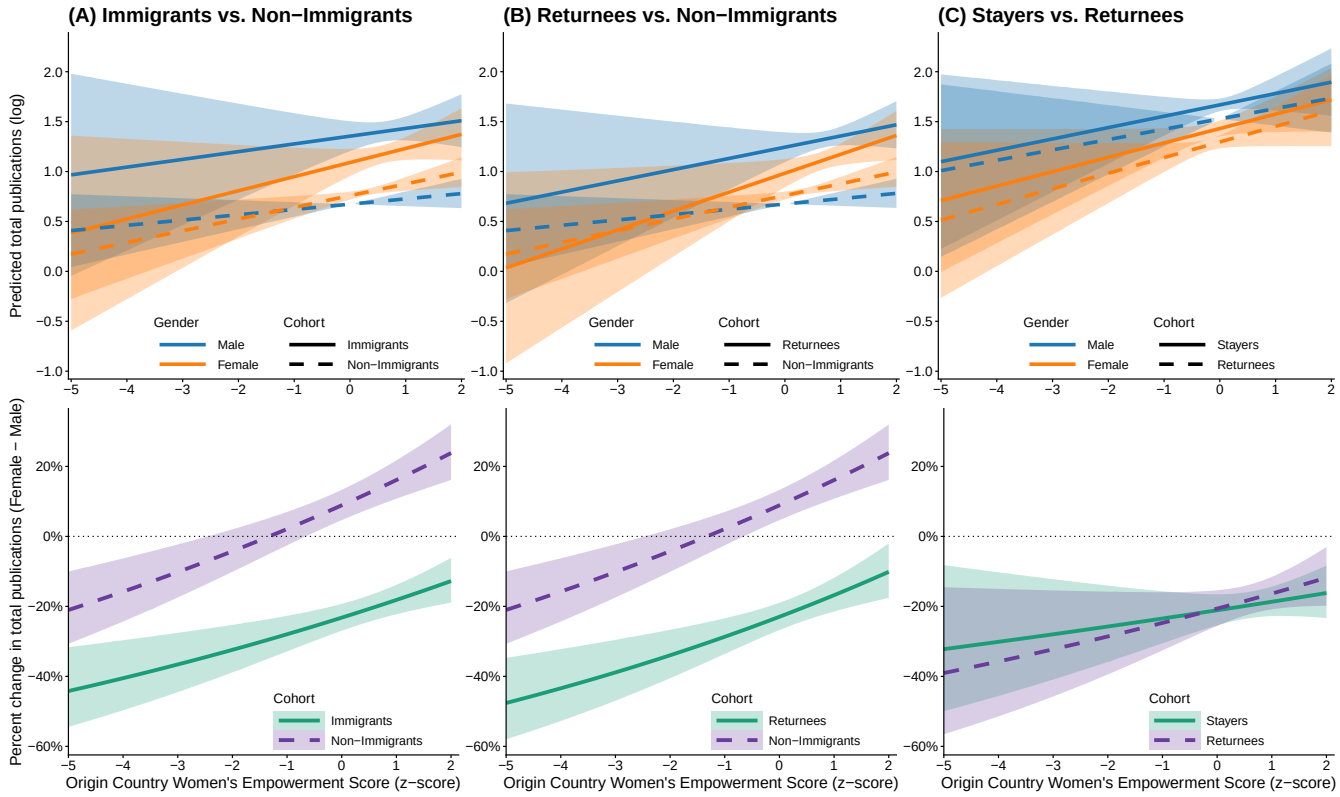


Fig. 4. Predicted total publications and gender gaps by women's empowerment across migration cohorts. The top row plots marginal predicted means of log total publications for men and women across the origin country women's empowerment (z-score); the bottom row shows the percent change in publications for women relative to men, with 95% confidence intervals. Panels correspond to cohort comparisons: **(A)** Immigrants vs. Non-Immigrants, **(B)** Returnees vs. Non-Immigrants, and **(C)** Stayers vs. Returnees. Estimates are based on the models in Table 3, which control for region, field, and residualized log GDP per capita.

returnees varies by the women's empowerment score of their countries of origin. In the countries with the lowest women's empowerment score in the sample, the gap is about 28% while it is only 9% in the countries with a high empowerment score. These differences highlight that decisions to stay or return redistribute a large share of high-output research activity.

Gender and women's empowerment then correspond to who produces that output within each cohort. In origin countries at the extreme low end of the observed women's-empowerment distribution, female scholars without immigration experiences publish about 21% fewer papers than their male counterparts from the same countries; at the high end of the observed empowerment distribution, they publish about 24% more. Among migrants, women also remain less productive than men across the observed empowerment range, but the gap shrinks as empowerment rises. In the stayer-versus-returnee comparison, female returnees publish about 29% fewer papers than male returnees in moderately low-empowerment contexts (around two standard deviations below the mean) and about 12% fewer in high-empowerment contexts (around two standard deviations above the mean). The corresponding female shortfall among stayers falls from about 26% to about 16%. Female stayers nonetheless outpublish female returnees throughout the distribution, although that gap is largest when origin-country empowerment is low (see Fig. 4C).

Extended cohort comparisons in the Supplementary Information show similar patterns when immigrants, returnees, stayers, and movers are modeled separately: international experience remains associated with higher publication counts than non-migrant careers, and higher women's empowerment is associated with smaller gender gaps in productivity (see SI, Table S16 and Figure S13).

As a robustness check, we replicate these analyses using *average yearly* publications up to 2022 (the coverage provided by the publication dataset (53)) rather than total counts. This measure normalizes output by the length of each researcher's publishing career and thus focuses on publication intensity rather than volume. The annualized-output models do not replicate every coefficient from the total-publication specification one-for-one, but they replicate general pattern: mobile researchers continue to out-publish non-immigrants, and stayers continue to out-publish returnees (see SI, Table S17 and Fig. S14). They also show that gender gaps narrow sharply as women's empowerment rises. In the annualized models, female disadvantages within migrant cohorts are substantial in low-empowerment contexts but approach zero at the high end of the empowerment distribution. Together, these results show that women's empowerment in the country of origin is linked not only to *who* returns, but also to the relative scientific productivity of mobile and non-mobile researchers, underscoring the role of gender-inclusive institutions in global knowledge production.

DRAFT

Table 3. Total publications and women's empowerment by migration cohort. OLS regressions estimate log total publications as a function of gender, women's empowerment (z-score), their interaction, cohort indicators, and residualized log GDP per capita. Columns compare Immigrants vs. Non-Immigrants, Returnees vs. Non-Immigrants, and Stayers vs. Returnees. All models include fixed effects for region and field, with standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: log total publications			
	Immigrants vs. Non-Immigrants	Returnees vs. Non-Immigrants	Stayers vs. Returnees
Female	0.085***	0.085***	-0.231***
	(0.020)	(0.020)	(0.033)
Women's Empowerment (z-score)	0.053	0.053	0.104
	(0.037)	(0.037)	(0.088)
Female×Women's Empowerment (z-score)	0.064***	0.064***	0.053*
	(0.013)	(0.013)	(0.030)
Immigrant	0.681***		
	(0.077)		
Female×Immigrant	-0.349***		
	(0.030)		
Female×Women's Empowerment (z-score)×Immigrant	-0.000		
	(0.020)		
Returnee		0.569***	
		(0.076)	
Female×Returnee		-0.346***	
		(0.032)	
Female×Women's Empowerment (z-score)×Returnee		0.013	
		(0.022)	
Stayer			0.140***
			(0.032)
Female×Stayer			-0.006
			(0.021)
Female×Women's Empowerment (z-score)×Stayer			-0.022
			(0.032)
GDP per Capita (log) (Residualized)	-0.089*	-0.089*	-0.127**
	(0.040)	(0.040)	(0.041)
Region FE	Yes	Yes	Yes
Field FE	Yes	Yes	Yes
R^2	0.037	0.035	0.082
Adjusted R^2	0.037	0.035	0.081
Number of Observations	26,577,392	26,441,039	159,946
Significance levels: $^*p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$			

Discussion

Nations experience "brain drain" when skilled citizens remain abroad—often resulting from a failure to repatriate skilled emigrants following education abroad. As the growing knowledge economy expands, the international competition for talent intensifies, raising incentives to understand what may attract highly skilled emigrants back to their countries of origin increases, especially in emerging economies. Despite women's growing share of this talent pool, relatively fewer work has examined the return migration of highly skilled female migrants, who on average show much lower return rates than their male counterparts.

Our results show that rates of return migration can differ for male and female emigrants from the same origin country, which highlights that countries may need different strategies to access this growing pool of skilled workers.

Bolstering women's empowerment domestically may be one option. Limited female participation in education, politics, and the economy (i.e., women's empowerment (54)) is associated with less return migration following education or work experience in the U.S. (see Fig. 1). Across origin countries and time, as women's empowerment rises, return becomes more attractive for all emigrants, but especially for women: the gap between male and female return migration narrows. With high return migration rates, migration of workers from emerging economies is less likely to lead to a one-way loss of talents but more like a circulation of talent that can benefit origin countries. These patterns are also robust to reweighting the sample by origin-country representativeness using external DS-visa benchmarks, suggesting that the main return-migration relationships are not driven by over- or underrepresentation of particular sending countries in the resume data.

Domestic women's empowerment is most strongly associated with return migration for high-skilled female migrants. In contrast, the PhD enrollment results point to a broader origin-country pattern: migrants from lower-empowerment countries are more likely to enroll in U.S. PhD programs, especially among those who first arrive as students. This pattern is not driven by higher PhD enrollment among women; in the student cohort, women enroll at lower average rates than men, while among those who first arrive as workers the average gender difference is small and not statistically distinguishable from zero. Overall, migrants from countries with greater women's empowerment are less likely to pursue a U.S. PhD, suggesting that strong academic and career options at home may reduce the appeal of U.S. doctoral training. Among migrants who enter the U.S. as workers, higher empowerment is more strongly associated with reduced male enrollment compared to female enrollment, sharply compressing gender gaps so that they approach zero and can slightly reverse at high empowerment.

Linking careers to publications highlights what is at stake. Immigrants, returnees, and stayers (i.e., migrants who remain in the U.S.) all publish more peer-reviewed academic articles than comparable non-migrants from the same origin country. This suggests that international mobility concentrates scientific productivity. Among migrant cohorts, women publish less than men, but this gap is largest for migrants from countries with low women's empowerment and shrinks as empowerment rises. Among non-migrants, women in lower-empowerment countries are less productive than male compatriots; however, women in high-empowerment countries match or even exceed male productivity. Among migrants, the gap narrows but typically does not disappear. Women's empowerment in their origin country is thus tied not only to who returns or pursues further study, but also to how effectively women translate experience abroad into scientific contributions.

These patterns point to women's representation in domestic institutions as an important correlate of whether outward mobility deepens or mitigates losses of human capital. Where laws, norms, and labor markets restrict women's opportunities, skilled women who leave are less likely to come back and, when they do, face larger productivity penalties relative to men. With higher levels of women's right in countries of origin, return migration rates of highly skilled workers is higher, gender gaps in productivity are smaller, and the benefits of international experience are more evenly shared. This highlights that policies that expand women's legal rights, political voice, and labor-market access—alongside investments in research infrastructure, transparent hiring and promotion, and support for early-career scientists—may therefore help shape repatriation of highly skilled emigrants. National policies that facilitate circular mobility and cross-border collaboration can further reduce the trade-off between sending talent abroad and retaining its benefits at home.

Several limitations in our study merit caution. Gender is inferred from names in LinkedIn profiles rather than self-

DRAFT

reported, and publication records capture only one dimension of innovative productivity, omitting entrepreneurship, teaching, clinical work, and policy roles. The analysis is observational, so unobserved selection into migration, return, or doctoral training may remain, including differences in family resources, academic ability, career preferences, professional networks, visa constraints, or country-specific policy changes that coincide with shifts in women's empowerment. In addition, resume data are not fully representative of all skilled migrants across origin countries, which is why we benchmark sample composition against external administrative sources and re-estimate the main return models using country-level post-stratification weights. We therefore assess the stability of the results across a series of robustness checks, including alternative model specifications, representativeness-weighted return regressions, measures and timing of women's empowerment, gender-coding thresholds, PhD outcome definitions, and extended publication models (see SI, Section S4). These checks show that the main qualitative patterns are stable across reasonable analytic choices. Taken together, the evidence suggests that greater women's empowerment may be associated with greater fairness domestically and with a more gender-balanced global circulation of talent.

Materials and Methods

Data sources. *Resume data (Revelio Labs).* We use individual-level resume data from Revelio Labs, which provide linked records on jobs, degrees, and basic demographics for workers worldwide. Each person has a unique identifier that connects employment spells (job title, employer, location, start and end dates, industry) and education spells (degree type, field, institution, timing). These histories allow us to reconstruct ordered career and education trajectories and to locate each spell by country (see SI, Section S1.1).

Women's empowerment and macro indicators (V-Dem and UNDP). To capture gender institutions in origin countries, we merge in country–year indices from the Varieties of Democracy (V-Dem) project. Our main measure is the Women's Political Empowerment Index, which summarizes women's civil liberties, civic participation, and political representation; we standardize it to have mean zero and unit standard deviation in our sample. In robustness checks, we replace this index with related V-Dem measures and with gender-equality indicators from the United Nations Development Programme, including the Gender Development Index, the Gender Inequality Index, and their underlying components (see SI, Sections S1.1.2–S1.1.3).

Publication data (SciSciNet). Scientific productivity is measured using SciSciNet, a large-scale integrated dataset of publications, authors, affiliations, fields, and citation-based indicators built from the Microsoft Academic Graph. SciSciNet provides publication histories and standardized affiliation countries for tens of millions of researchers and has been widely used in studies of scientific mobility and careers (see SI, Section S1.1.4).

Sample construction and migration variables. We begin with all individuals in Revelio whose first recorded spell occurs outside the U.S. and who later have at least one spell in the United States. We drop records with missing or invalid start dates, restrict to years up to 2019, and keep only those whose highest degree is at least a bachelor's. Individuals with unknown gender are excluded.

For each person, we sort all employment and education spells chronologically to form a single timeline. We define the country of the first spell as the country of origin, and the first year with a U.S. spell as the immigration year. Entry route is classified as *student* when the first U.S. spell is education and *worker* when it is employment. Total time spent in the United States is computed by aggregating the duration of U.S. spells.

Migration outcomes are defined from the last recorded spell: we label individuals as *stayers* if the last spell is in the United States, *returnees* if it is in the country of origin, and *movers* if it is in a third country. A binary return indicator contrasts returnees with all others, and the last-action year is taken as the start year of the final spell. Within U.S. experience, we record whether individuals have ever held a U.S. job, ever studied in the United States, and ever enrolled in a U.S. PhD program. The final analytical sample contains 2,620,865 skilled migrants, of whom 1,092,010 first arrive as students and 1,528,855 as workers (see SI, Sections S1.2 and S2).

Women's empowerment timing and linkage. V-Dem and UNDP indicators are merged to each individual by origin country and year using two timing conventions. In our baseline, women's empowerment is measured in the immigration year, capturing conditions at departure. A robustness specification instead measures empowerment in the last-action year, capturing conditions at the time of the observed return, stay, or onward move. All indices are standardized before merging, so coefficients can be interpreted in terms of standard-deviation changes (see SI, Sections S1.1.2, S1.1.3, and S4.4).

Linking migrants to publications and defining comparison groups. To connect migrants' careers to their scientific output, we link Revelio individuals to SciSciNet authors via names and affiliations. We first harmonize names in both datasets (lowercasing, removing accents and punctuation, splitting into first, middle, and last names) and generate candidate links using exact matches and rules based on compatible initials. We then disambiguate these candidates with affiliation histories: institution names are normalized and mapped to canonical identifiers, candidate pairs must share at least one institution, and when multiple matches remain we retain the pair with the greatest overlap. Revelio profiles that still match more than one SciSciNet author are dropped to enforce one-to-one links (see SI, Section S1.1.4).

Linked individuals form our group of *immigrant authors*. We define *non-immigrant authors* as SciSciNet researchers whose entire affiliation history lies within a single country and who are not linked to any immigrant. Across both groups, we observe 206,948 immigrant authors and 29,942,429 non-immigrant authors (see SI, Table S5).

For each author, we construct two outcome measures: total publications up to 2022, and average yearly publications, defined as total publications divided by career length (2022 minus the year of first publication). Both outcomes are used in log form in the regression analysis because of strong right skew (see SI, Figures S11–S12).

Representativeness and external benchmarks. Public administrative data on skilled migrants do not track full careers but provide benchmarks for the coverage of our resume sample. We therefore compare the size and composition of our immigrant sample to two sources.

First, we use annual counts of F-1 student visas and H-1B specialty-occupation visas issued by the U.S. Department of State from 1997 to 2019. These statistics record visa issuances and petition approvals rather than distinct people: the same individual can receive multiple F-1 visas, and many H-1B approvals extend or modify the status of workers already in the United States. We treat these series as upper bounds on the number of potential migrants in each cohort and interpret our coverage ratios (resume counts divided by visa counts) as conservative lower bounds on the share of skilled immigrants observed.

Second, for students, we use the Department of Homeland Security's SEVIS "SEVIS by the Numbers" reports for 2014–2019, which provide counts of active international-student records by gender. These snapshots measure a stock of active records at a given time, including continuing students and those on post-completion training, and may contain multiple records per person. They therefore also overstate the number of distinct individuals but offer the only public source with gender-disaggregated information on international students. Figures S8–S10 and Table S6 in the SI show that the resume-based immigrant counts track the rise and fall of skilled inflows over time, remain below administrative totals as expected, and achieve relatively higher coverage for English-

DRAFT

speaking and high-income origin countries than for large senders such as China and India. In an additional robustness check, we use these same country-level DS-visa shares to construct post-stratification weights, defined as the ratio of each origin country's share in DS-visa records to its share in the analytical sample, and re-estimate the baseline return-migration regressions under this weighting scheme.

Regression specifications and robustness checks. We analyze three families of outcomes: return migration, enrollment in a U.S. PhD program, and scientific productivity.

For return migration and PhD enrollment, we estimate logistic regressions separately for students and workers. Return models use a binary indicator for whether an individual ultimately returns to their country of origin; PhD models use a binary indicator for whether they ever enroll in a U.S. PhD program. In all specifications, the key regressors are gender, women's empowerment, and their interaction. Controls include home-country income (entered as residualized or raw log GDP per capita in different specifications), time spent in the United States, and indicators for U.S. work and education experience, along with fixed effects for year and origin region. Standard errors are clustered by country of origin. We test robustness to alternative GDP controls, alternative empowerment measures from V-Dem and UNDP, alternative timing of the empowerment index (immigration vs. last-action year), alternative thresholds used to define the female indicator from gender probabilities, and country-level post-stratification weights based on the ratio of DS-visa shares to sample shares by origin country (see SI, Section S4). This weighting exercise assesses whether the main return-migration results are sensitive to imbalances in the origin-country composition of the resume sample relative to external administrative benchmarks.

For publication outcomes, we estimate ordinary least squares models with log total publications or log average yearly publications as the dependent variable. We include indicators for migration cohort (immigrants, returnees, stayers, movers, non-immigrants), gender, women's empowerment, and their interactions, together with home-country income and fixed effects for region and field; standard errors are clustered by country of origin. From these models, we compute marginal predicted means for men and women along the women's-empowerment distribution and translate differences in log outcomes into percentage gaps. The corresponding prediction curves and gender gaps by cohort are visualized in the SI (Figures S13–S14). Across these specifications and robustness checks, the broad patterns are consistent: internationally mobile researchers tend to out-publish non-immigrants, stayers tend to out-publish returnees, and higher women's empowerment is associated with smaller gender gaps in scientific productivity, even though the magnitude of these relationships differs between cumulative and annualized publication measures.

ACKNOWLEDGMENTS. This work is supported by the University of Pittsburgh Center for Research Computing. We are grateful for input from Qing Chang.

1. K Hoisl, Tracing mobile inventors—the causality between inventor mobility and inventor productivity. *Res. policy* **36**, 619–636 (2007).
2. L Ackers, Internationalisation, mobility and metrics: A new form of indirect discrimination? *Minerva* **46**, 411–435 (2008).
3. A Fernandez-Zubieta, A Geuna, C Lawson, Researchers' mobility and its impact on scientific productivity. *Univ. Turin working paper* (2013).
4. C Franzoni, G Scellato, P Stephan, The mover's advantage: The superior performance of migrant scientists. *Econ. Lett.* **122**, 89–93 (2014).
5. P Azoulay, I Ganguli, JG Zivin, The mobility of elite life scientists: Professional and personal determinants. *Res. policy* **46**, 573–590 (2017).
6. N Netz, S Hampel, V Aman, What effects does international mobility have on scientists' careers? a systematic review. *Res. evaluation* **29**, 327–351 (2020).
7. B Ghosh, *Migrants' remittances and development: myths, rhetoric and realities.* (International Organization for Migration (IOM)), (2006).
8. B Klagge, K Klein-Hitpaß, High-skilled return migration and knowledge-based economic development in regional perspective: Conceptual considerations and the example of poland, (CMR working papers), Technical report (2007).
9. SP Kerr, W Kerr, Ç Özden, C Parsons, Global talent flows. *J. Econ. Perspectives* **30**, 83–106 (2016).
10. SP Kerr, W Kerr, Ç Özden, C Parsons, High-skilled migration and agglomeration. *Annu. Rev. Econ.* **9**, 201–234 (2017).
11. Z Koczan, G Peri, M Pinat, D Rozhkov, *The impact of international migration on inclusive growth: A review.* (International Monetary Fund), (2021).
12. S Bernstein, R Diamond, A Jiranaphawiboon, T McQuade, B Pousada, The contribution of high-skilled immigrants to innovation in the united states, (National Bureau of Economic Research), Technical report (2022).
13. J Hunt, M Gauthier-Loiselle, How much does immigration boost innovation? *Am. Econ. Journal: Macroecon.* **2**, 31–56 (2010).
14. WR Kerr, WF Lincoln, The supply side of innovation: H-1b visa reforms and us ethnic invention. *J. Labor Econ.* **28**, 473–508 (2010).
15. G Chellaraj, KE Maskus, A Mattoo, The contribution of international graduate students to us innovation. *Rev. Int. Econ.* **16**, 444–462 (2008).
16. National Science Board, Science and engineering indicators 2014 (2014) Online; accessed 2014.
17. S Anderson, Immigrants and billion-dollar companies. *NFAP Policy Brief* (2018).
18. JV Beaverstock, Highly skilled international labour migration and world cities: Expatriates, executives and entrepreneurs in *International handbook of globalization and world cities.* (Edward Elgar Publishing), (2011).
19. BC Holding, C Acciai, JW Schneider, MW Nielsen, Quantifying the mover's advantage: transatlantic migration, employment prestige, and scientific performance. *High. Educ.* **87**, 1749–1767 (2024).
20. UNESCO, *UNESCO Science Report: Towards 2030.* (UNESCO Publishing, Paris), (2015).
21. H Boekhout, I van der Weijden, L Waltman, Gender differences in scientific careers: A large-scale bibliometric analysis. *arXiv preprint arXiv:2106.12624* (2021).
22. U.S. Department of Commerce, Women in stem representation matters (2024) Accessed June 22, 2026.
23. X Zhao, A Akbaritabar, R Kashyap, E Zagheni, A gender perspective on the global migration of scholars. *Proc. Natl. Acad. Sci.* **120**, e2214664120 (2023).
24. M Kwiek, L Szymula, Young male and female scientists: A quantitative exploratory study of the changing demographics of the global scientific workforce. *Quant. Sci. Stud.* **4**, 902–937 (2023).
25. SC Pearce, Today's immigrant woman entrepreneur. *The Divers. Factor* **13**, 23–29 (2005).
26. WW Ding, F Murray, TE Stuart, From bench to board: Gender differences in university scientists' participation in commercial science (2010).
27. W Shi, R Wu, Women's strength in science: exploring the influence of female participation on research impact and innovation. *Scientometrics* **129**, 4529–4551 (2024).
28. S Donato, C Corradi, U Di Maggio, Rethinking women's return migration: Evidence from tunisia and morocco. *Societies* **15**, 180 (2025).
29. N Auraiampai, D Ayuwat, A Jampaklay, F Onaka, Entrepreneurial processes of thai female returning migrants. *Comp. Sociol.* **22**, 717–739 (2023).
30. A George, "Return Migration and Entrepreneurship: The Role of Highly Skilled Women in Nigeria's Technology Services Sector," PhD thesis, Open University (United Kingdom) (2023).
31. C Fry, JL Furman, Migration and global network formation: Evidence from female scientists in developing countries. *Organ. Sci.* **36**, 1047–1071 (2025).
32. A Anastasiadou, E Zagheni, HAG de Valk, Including the gender dimension of migration is essential to avoid systematic bias in migration predictions. *Proc. Natl. Acad. Sci.* **122**, e2500874122 (2025).
33. A Anastasiadou, J Kim, E Sanlitürk, HAG de Valk, E Zagheni, Gender differences in the migration process: A narrative literature review. *Popul. Dev. Rev.* **50**, 961–996 (2024).
34. F Docquier, BL Lowell, A Marfouk, A gendered assessment of highly skilled emigration. *Popul. Dev. review* **35**, 297–321 (2009).
35. S Donato, Women migrant returnees as intermediaries: Exploring empowerment and agency of migrant women returnees in the eu-mena region. *J. Int. Migr. Integration* pp. 1–17 (2023).
36. J Lewis, M Swannell, The macroeconomic determinants of migration. *SSRN Electron. J.* (2018).
37. A Kwilinski, et al., International migration drivers: Economic, environmental, social, and political effects. *Sustainability* **14**, 6413 (2022).
38. JD Angrist, JS Pischke, *Mostly harmless econometrics: An empiricist's companion.* (Princeton university press), (2009).
39. M Piracha, Occupational choice of return migrants. *IZA World Labor* **197**, 1–10 (2015).
40. T Gluszynski, V Peters, *Survey of earned doctorates: A profile of doctoral degree recipients.* (Statistics Canada, Culture, Tourism and the Centre for Education Statistics), (2005).
41. M Roach, H Sauermann, A taste for science? phd scientists' academic orientation and self-selection into research careers in industry. *Res. policy* **39**, 422–434 (2010).
42. M Roach, Encouraging entrepreneurship in university labs: Research activities, research outputs, and early doctorate careers. *PloS one* **12**, e0170444 (2017).
43. MG Finn, LA Pennington, Stay rates of foreign doctorate recipients from u.s. universities, 2013, (Oak Ridge Institute for Science and Education), Report (2018).
44. National Science Board, Science and engineering indicators 2018, (National Science Foundation, National Center for Science and Engineering Statistics (NCSES), Alexandria, VA), Technical Report NSB-2018-1 (2018).
45. RB Freeman, What does global expansion of higher education mean for the united states? in *American Universities in a global market.* (University of Chicago Press), pp. 373–404 (2010).
46. ES Brezis, Why migrate: For study or for work? *Economies* **4**, 17 (2016).
47. JI Carrasco, A Akbaritabar, M Godin, C Vargas-Silva, Prioritizing global equity in migration research. *Humanit. Soc. Sci. Commun.* **13**, 16 (2025).
48. A Rottenkolber, et al., It's who you know—unless you're famous: professional networks and prestige in scholarly mobility, (Max Planck Institute for Demographic Research, Rostock, Germany), Technical report (2025).
49. H Li, CJ Tessone, A Zeng, Productive scientists are associated with lower disruption in scientific publishing. *Proc. Natl. Acad. Sci.* **121**, e2322462121 (2024).
50. AM Petersen, FJ Arroyave, F Pammolli, The disruption index suffers from citation inflation: Re-analysis of temporal cd trend and relationship with team size reveal discrepancies. *J. Informetrics* **19**, 101605 (2025).

DRAFT

51. N Deng, X Gu, Y Fan, S Havlin, A Zeng, The critical role of persistent disruption in advancing science. *Nat. Comput. Sci.* **5**, 1–10 (2025).
52. H Jiang, J Zhou, Y Ding, A Zeng, Overcoming recognition delays in disruptive research: The impact of team size, familiarity, and reputation. *J. Informetrics* **18**, 101549 (2024).
53. Z Lin, Y Yin, L Liu, D Wang, Sciscinet: A large-scale open data lake for the science of science research. *Sci. Data* **10**, 315 (2023).
54. VM Moghadam, What is democracy? promises and perils of the arab spring. *Curr. Sociol.* **61**, 393–408 (2013).
55. D Weitzel, J Gerring, D Pemstein, SE Skaaning, Measuring electoral democracy with observables. *Am. J. Polit. Sci.* (2023).
56. L Lott, J Spannagel, Quality assessment of the academic freedom index: Strengths, weaknesses, and how best to use it. *Perspectives on Polit.* p. 1–23 (2025).
57. F Jiang, T Pan, J Wang, Y Ma, To academia or industry: Mobility and impact on acm fellows' scientific careers. *Inf. Process. & Manag.* **61**, 103736 (2024).
58. Y Yang, et al., Gender differences in dropout rate: From field, career status, and generation perspectives. *J. Informetrics* **19**, 101615 (2025).
59. S Zhang, F Liu, H Xia, Quantifying spatial–temporal citation diffusion of individual papers in knowledge space. *J. Assoc. for Inf. Sci. Technol.* **76**, 1682–1700 (2025).

DRAFT

Supporting Information

S1. Materials and Methods

S1.1. Datasets.

S1.1.1. Resume Dataset. The dataset from Revelio Labs[‡] provides structured, individual-level employment and education records that we use to reconstruct skilled workers' career and migration trajectories. It links job spells and degrees for each person, allowing us to follow movements across countries, sectors, and educational systems in a way that is well suited to studying brain drain, brain gain, and return migration.

We rely on three core components:

- **Position:** Contains job-specific information, including job title, employer name, start and end dates, and industry classification codes. These records allow us to identify work episodes inside and outside the United States, measure duration in each location, and distinguish workers by sector.
- **Education:** Documents formal degrees, including degree type, field of study, institution, and timing. These records are used to classify highest educational attainment, identify U.S. versus non-U.S. training, and distinguish those who first arrive as students.
- **User:** Provides individual-level identifiers and demographics, including names, gender[§], and a unique user ID that links Position and Education records for the same person.

Combining these components yields longitudinal profiles in which each row in our analytical dataset corresponds to a unique individual whose employment and education spells are ordered over time and geocoded by country. This structure allows us to classify migrants by entry route (student vs. worker), determine whether and when they return to their country of origin, and analyze gender differences in return, retention, and onward mobility among highly skilled professionals.

S1.1.2. Varieties of Democracy (V-Dem). The *Varieties of Democracy (V-Dem)* project[¶] provides annual, country–year data for more than 200 countries and territories, with a wide set of indicators covering electoral democracy, liberal and participatory institutions, civil liberties, and patterns of inclusion and exclusion. The data are based on multiple expert coders per country–year and estimated using a Bayesian item–response framework that corrects for coder disagreement and differential reliability, yielding comparable indices over time and across countries.[‖] V-Dem is now standard in comparative politics and political economy for measuring institutional quality, democratic backsliding, and the distribution of power (55, 56).

Our main measure of women's empowerment is the V-Dem *Women's Political Empowerment Index* (variable `v2x_gender`). This index summarizes the extent to which women enjoy civil liberties, participate in civil society, and are represented in politics and other decision-making bodies, and is defined by V-Dem as capturing "a process of increasing capacity for women, leading to greater choice, agency, and participation in societal decision-making." In the analyses, we transform this index into a standardized z-score within our sample so that a one-unit change corresponds to a one–standard deviation difference in women's political empowerment.

In robustness checks (Tables S9 and S10), we replace this composite index with several related V-Dem indicators from the same family of measures. These include component indices for women's civil liberties (`v2x_gencl`), women's civil society participation (`v2x_gencs`), and women's political participation and representation (`v2x_genpp`), each aggregating multiple underlying items on freedoms, organizational life, and access to elected office. All of these variables are re-scaled to z-scores and merged to individuals by country of origin and year, providing a set of conceptually related but distinct measures of the gender environment in which migrants make education, career, and return decisions.

S1.1.3. United Nations Development Programme (UNDP). The United Nations Development Programme (UNDP) publishes the *Human Development Report* and an accompanying data series that provide harmonized, country–year measures of human development and gender inequality.[**] These data cover most countries from the early 1990s onward and report indicators on income, health, education, and gender gaps in these domains.

For this study, we draw on UNDP's gender–related human–development indicators as alternative measures of conditions facing women in the country of origin. In particular, the first blocks of Tables S9 and S10 use UNDP variables such as the *Gender Development Index (GDI)*, which captures gaps between women and men in health, education, and living standards, and the *Gender Inequality Index (GII)*, which summarizes inequality in reproductive health, political representation, educational attainment, and labour–force participation.[††] We also use the underlying female and male indicators that feed into these indices—for example, female labour–force participation rates and women's share of parliamentary seats—as robustness checks for the baseline women's–empowerment measure.

All UNDP indicators are merged to individuals' records at the country–of–origin and year level, aligning the timing of the gender–development environment with migrants' immigration or decision years (as described in Section S1.2). This integration allows us to test whether the patterns documented in the main text hold when women's empowerment is operationalized using established UNDP gender–equality metrics.

S1.1.4. SciSciNet. *SciSciNet* is an openly released, large-scale "data lake" designed for quantitative science-of-science research (53). Built from the final Microsoft Academic Graph snapshot,[‡‡] it provides a harmonized corpus of scholarly publications with deduplicated records and standardized identifiers for authors, affiliations, venues, fields, and document types. The core bibliographic layer is complemented by rich linkages to external data sources, including public-sector grants, patents, clinical trials, news coverage, and social-media mentions, enabling analyses of how knowledge is produced, funded, and diffused across sectors and over time.

Beyond raw metadata, SciSciNet includes a suite of pre-computed indicators at the paper and author level—such as multi-year citation counts, disruption and novelty scores, citation-trajectory parameters, and demographic and career attributes—to facilitate large-scale empirical work. The full dataset, processing pipeline, and documentation are hosted on Figshare and GitHub, supporting transparent validation and community-driven extensions (53). SciSciNet has already been used to study how institutional mobility shapes elite scientists' productivity and citation impact (57), document gendered differences in career persistence and dropout across fields and

[‡] https://www.reveliolabs.com/

[§] https://www.data-dictionary.reveliolabs.com/methodology.html#gender-and-ethnicity

[¶] https://www.v-dem.net/

[‖] vdem_codebook, vdem_methodology

[**] https://hdr.undp.org

[††] For definitions of the GDI and GII, see the UNDP Human Development Data Center: https://hdr.undp.org/data-center.

[‡‡] https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/

DRAFT

generations (58), quantify how citations diffuse over time through an embedded "knowledge space" (59), and analyze persistently disruptive contributions and their relationship to team size (51).

In this study, we use SciSciNet to measure scientific productivity for the individuals linked to the Revelio dataset (see Section S1.2), and to construct a comparison group of non-immigrant researchers. Linking author-level publication records to migrants' career histories allows us to quantify total and average yearly publication output, compare immigrants, returnees, stayers, movers, and non-immigrants, and relate these outcomes to home-country women's empowerment and income. Together, Revelio, V-Dem, UNDP, and SciSciNet provide an integrated framework for analyzing how skilled migration and gendered institutional environments shape both the geography and volume of scientific production.

S1.2. Data Preparation and Construction of the Analytical Dataset. Our analytical dataset is built from the *Revelio Labs* resume data (Section S1.1.1), which provide longitudinal employment and education records at the individual level. We clean and harmonize these records, construct migration and demographic variables, and then merge in country–year indicators from V-Dem and UNDP (Sections S1.1.2–S1.1.3) and publication outcomes from SciSciNet (Section S1.1.4). The final structure is one row per individual summarizing their origin, migration history, and U.S. experiences, together with linked institutional and productivity measures.

S1.2.1. Data Cleaning and Preprocessing. The preprocessing pipeline proceeds in three main steps.

- **Filtering and missing data handling.**
  - We drop users whose employment or education records lack valid start dates.
  - We exclude users with unknown gender, as gender is central to all analyses.
- **Standardizing education and employment spells.**
  - For each record, we correct misordered start and end dates where possible.
  - We restrict the sample to records up to 2019 to focus on pre-pandemic migration patterns.
  - We retain only individuals who hold at least one of the following degrees: *Bachelor, Master, MBA, or Doctorate*.
- **Defining the migrant sample and trajectories.**
  - We restrict to individuals whose first recorded experience occurs *outside* the United States and who later have at least one record in the United States.
  - For each user, we sort all employment and education records chronologically and assign an order index, reconstructing a unified career and education timeline.

After these steps, we obtain a cohort of internationally mobile, highly educated individuals whose movements into and out of the United States can be traced over time.

S1.2.2. Variable Construction. From the cleaned and ordered records, we construct the following key variables used in the analyses.

User identification and demographics

- `user_id`: Unique identifier for each individual.
- `user_name`: First name of the individual.
- `user_surname`: Surname of the individual.
- `first_record_year`: Year of the individual's first recorded employment or education spell.
- `is_female`: Indicator equal to `True` if the inferred probability of being female is at least 0.5.
- `country_of_origin`: Country associated with the first recorded experience.
- `country_of_origin_iso3c`: ISO 3166-1 alpha-3 country code of the country of origin.
- `country_of_origin_region`: Geographic region corresponding to the country of origin.

Migration and return outcomes

- `immigration_year`: First year in which the individual has an employment or education record in the United States.
- `first_is_edu`: Indicator equal to 1 if the first record in the United States is an education spell (used to classify *students* vs. *workers*).
- `years_in_us`: Total duration spent in the United States, aggregating employment and education spells.
- `last_action_country`: Country associated with the individual's most recent record.
- `last_action`: Categorical outcome summarizing migration status at the end of the observed trajectory: *Returned* (last record in country of origin), *Stayed* (last record in the United States), or *Moved* (last record in a third country).
- `returned`: Boolean indicating whether the individual returned to their country of origin (*Returned* vs. all other outcomes).
- `last_action_year`: Year corresponding to the start date of the last employment or education spell.

Employment and education history in the United States

- `us_job_before_last_action`: Indicator for whether the individual held at least one job in the United States prior to their last recorded action.
- `us_edu_before_last_action`: Indicator for whether the individual pursued at least one education spell in the United States prior to their last recorded action.
- `doctor_in_us`: Indicator equal to 1 if the individual enrolled in a U.S. PhD program at any point in their career.

After applying these definitions, the final analytical sample includes **2,620,865** individuals. Among them, **1,092,010** are classified as *students* (their first record in the United States is an education spell, `first_is_edu = 1`), and **1,528,855** are classified as *workers* (their first record in the United States is an employment spell, `first_is_edu = 0`).

DRAFT

S1.2.3. Integration of V-Dem Data. To incorporate institutional and governance characteristics, we merge the V-Dem country–year data (Section S1.1.2) to the individual records. The merge uses `country_of_origin_iso3c` and either `immigration_year` or `last_action_year`, depending on the timing convention:

- **Immigration year ("First"):** women's empowerment is measured in the year the individual first appears in the United States, capturing the baseline gender environment at departure.
- **Decision year ("Last"):** women's empowerment is measured in the year of the last observed record, capturing the institutional environment at the time the return/stay/move decision is made.

In both cases, the V-Dem Women's Political Empowerment Index and related indicators are transformed to z-scores within our sample. This integration allows us to link each migrant's trajectory to the gender and institutional environment in their country of origin at relevant decision points.

S1.2.4. Linking SciSciNet Publication Records to Immigrants' Career Histories. We next link author-level publication data from **SciSciNet** (Section S1.1.4) to the career histories derived from the Revelio dataset. Because the two sources do not share persistent identifiers, the linkage proceeds in two stages: deterministic name matching followed by affiliation-based disambiguation.

Deterministic name matching We first harmonize names in both datasets by lowercasing, converting to ASCII, removing punctuation, and splitting into *first*, *middle* (possibly empty), and *last* name. We then apply the following rules, in order:

1. **Exact rule:** Accept a candidate pair when the cleaned full names are identical.
2. **Initial-compatible rule:** If the exact match fails, accept a pair when all of the following hold:
   (a) last names match verbatim;
   (b) first names either match verbatim or one side is a single-letter initial equal to the first letter of the other;
   (c) middle names are either absent on at least one side or consist of single tokens that satisfy the same initial-letter logic (multi-word middle tokens are excluded).

These rules yield a set of candidate links between SciSciNet authors and Revelio users that are plausible on the basis of names alone.

Affiliation-based disambiguation Name-based matches can still map one SciSciNet author to multiple Revelio profiles (and vice versa). To enforce a strict one-to-one mapping, we use affiliation information recorded in both datasets:

1. **Affiliation normalization:** We lower-case organization names, remove punctuation and diacritics, and map them to canonical institution keys using the Research Organization Registry (ROR).[§§]
2. **Overlap scoring:** For each candidate author–user pair, we count the number of shared canonical institutions across their affiliation histories and discard pairs with zero overlap.
3. **Best-hit selection:** If a SciSciNet author matches multiple Revelio profiles, we retain only the pair with the highest overlap count.
4. **Uniqueness of Revelio IDs:** We remove any Revelio profile that still matches more than one SciSciNet author after the previous filters. The remaining pairs form a unique bridge between the two datasets.

This procedure yields two mutually exclusive groups:

- **Immigrants:** Revelio-identified migrants who are successfully linked to a SciSciNet author.
- **Non-immigrants:** SciSciNet authors whose full publication-affiliation history lies within a single country (based on ISO-3 codes) and who do not appear in the immigrant group.

Across the linked SciSciNet–Revelio corpus, we observe 206,948 immigrant authors and 29,942,429 non-immigrant authors, for a total of 30,149,377 unique researchers (see Table S5 for the breakdown by country of origin and gender). These linked data underpin the analyses of scientific productivity and its relationship to migration status and women's empowerment.

S2. Descriptive Patterns by Origin, Gender, and Status

This section provides an overview of the composition of the skilled immigrant sample and the linked author dataset, focusing on origin regions and countries, gender, entry channel, geographic distribution within the United States, work and study histories, and educational attainment.

Table S1 summarizes the number of immigrants by World Bank region of origin, disaggregated into female and male workers and students. Europe & Central Asia and East Asia & Pacific account for the largest shares of skilled immigrants in the data, followed by South Asia and Latin America & Caribbean. Within each region, the table distinguishes whether individuals first enter the United States as workers or as students, and shows that both channels are present across all regions, with notable gender differences in the relative size of worker and student streams.

Figure S1 tracks the evolution of skilled immigration flows over time by gender from 1990 to 2019 (plotted on a logarithmic scale). Both male and female immigrant counts rise over the period, with particularly strong growth in the late 1990s and 2000s. Male flows remain somewhat higher in absolute terms, but the figure shows that female skilled immigration has expanded in parallel, indicating a growing female presence in high-skill migration to the United States.

Table S2 provides a more granular view by listing immigrant counts for the top origin countries, again broken down by gender and entry type. India, the United Kingdom, China, and Canada are the four largest sending countries in this skilled sample, with substantial numbers of both female and male workers and students. The table highlights cross-country variation in gender composition and in the balance between work-based and study-based entry, which motivates later analyses that condition on country of origin and entry cohort.

Figure S2 examines last observed migration status (Stayed, Returned, Moved onward) for the top 20 origin countries, separately for women and men. Each stacked bar shows the share of skilled immigrants who remain in the United States, return to their country of origin, or move to a third country. The figure reveals substantial cross-country heterogeneity in these outcomes and systematic gender differences within countries, underscoring that return and onward mobility are shaped jointly by origin context and gender.

[§§] https://ror.org/

DRAFT

Figure S3 shows the distribution of the number of years that immigrants are observed residing in the United States based on their resume records. Most individuals are present for a moderate duration, with a long right tail capturing a smaller group with extended U.S. stays. This distribution provides context for interpreting differences in return behavior by tenure in the United States.

Table S3 reports the distribution of immigrants by their last recorded U.S. state and gender. A substantial share of records have an unknown state, but among identifiable locations, California, New York, Illinois, Texas, Florida, and Massachusetts account for the largest immigrant populations. The table documents both the concentration of skilled migrants in a handful of major destination states and the broad geographic spread across the rest of the country.

Figure S4 distinguishes whether individuals' first recorded experience in the United States is as a worker or as a student. A majority (58%) first appear in the data through an education record, while 42% enter via employment. This split justifies the later distinction between "students" and "workers" as analytically meaningful entry cohorts in the regression analysis.

Figure S5 summarizes whether immigrants held at least one U.S. job before their last observed migration outcome (Stayed, Returned, or Moved). Among those who stay in the United States, 77% have prior U.S. work experience. For returnees, this share is 51%, while for onward movers it is 72%. These patterns indicate that a majority of skilled migrants—particularly stayers and onward movers—accumulate U.S. employment experience before their trajectory is finalized.

Figure S6 restricts the same measure to the student cohort, i.e., individuals whose first U.S. record is education. Among student stayers, 64% held at least one U.S. job before their last action; for student returnees, 49% did; and for student onward movers, 64% did. Thus, even among those who enter through education, a substantial fraction gain work experience in the U.S. before returning or moving on.

Figure S7 displays the distribution of the number of distinct jobs held in the United States, separately for women and men (with the x-axis truncated at 10 jobs for readability; the true maximum is 20). The distribution is highly right-skewed: most immigrants hold only a small number of jobs, while a relatively small group experiences many job changes. The gender-specific histograms are broadly similar, indicating comparable patterns of job mobility within the United States for female and male skilled migrants.

Table S4 reports the highest degree obtained by immigrants, by gender. For both women and men, Bachelor's and Master's degrees are the most common, with roughly 40–45% holding a Bachelor's degree and about one-third holding a Master's degree. MBAs and Doctorates are less frequent but still substantial, particularly among men. This distribution confirms that the sample is highly educated and that advanced degrees are common among skilled migrants in the data.

Finally, Table S5 summarizes the linked author dataset used in the publication analysis, listing the number of immigrant and non-immigrant authors by country of origin and gender, along with the ratio of immigrant to non-immigrant authors. Large sending countries such as China, India, the United Kingdom, and Germany contribute both substantial immigrant author populations and very large non-immigrant baselines. The immigrant-to-non-immigrant ratios vary widely across countries, reflecting differences in how strongly national scientific communities are represented among internationally mobile researchers.

Table S1. Immigrants by region of origin, gender, and entry cohort. Counts of skilled immigrants in the analytical sample by World Bank region of origin, showing the total number of immigrants and their distribution across workers and students, disaggregated by gender. Workers are individuals whose first U.S. record is employment; students are those whose first U.S. record is education.

Region of Origin		Workers		Students	
Name	Total	Female	Male	Female	Male
Europe & Central Asia	775,298	214,462	259,768	135,907	165,161
East Asia & Pacific	601,786	164,217	178,090	118,661	140,818
South Asia	441,475	91,406	159,463	63,992	126,614
Latin America & Caribbean	373,948	95,049	116,443	68,277	94,179
North America	205,822	59,804	73,300	35,577	37,141
Middle East & North Africa	149,461	28,001	47,232	24,712	49,516
Sub-Saharan Africa	71,294	16,910	24,200	12,169	18,015

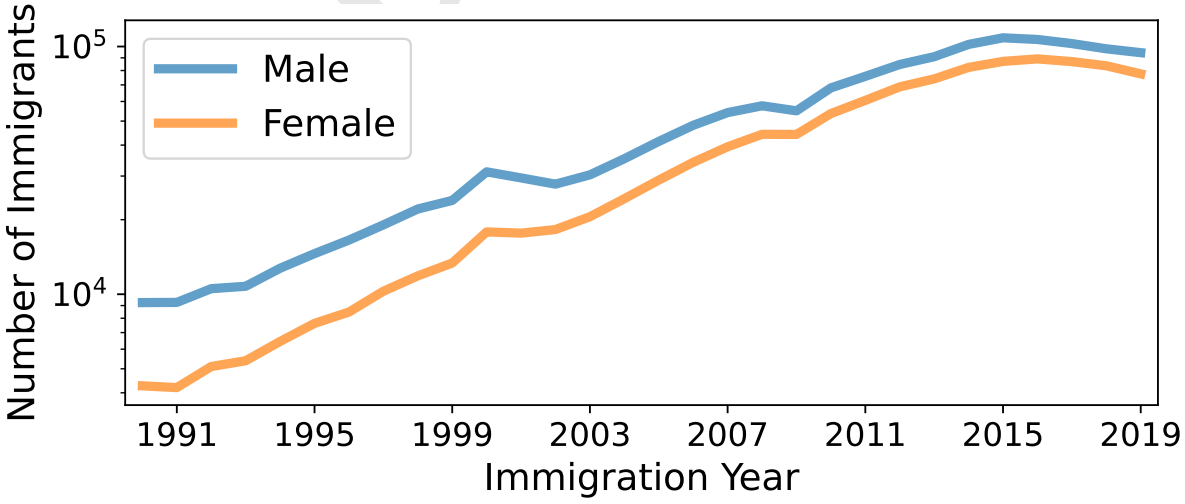


Fig. S1. Immigration to the United States by gender, 1990–2019. Yearly counts of male and female immigrants in the Revelio resume data, plotted on a logarithmic scale for the number of immigrants.

DRAFT

Table S2. Immigrants by country of origin, gender, and entry cohort. Counts of skilled immigrants in the analytical sample by country of origin (name and ISO3C code), reporting the total number of immigrants and their distribution across workers and students, disaggregated by gender. Workers are individuals whose first U.S. record is employment; students are those whose first U.S. record is education.

Country of Origin			Workers		Students	
Name	ISO3C	Total	Female	Male	Female	Male
India	IND	392,428	82,356	140,056	57,590	112,426
United Kingdom	GBR	238,917	68,875	94,672	34,401	40,969
China	CHN	210,121	53,795	62,057	42,308	51,961
Canada	CAN	204,434	59,635	73,162	34,988	36,649
Philippines	PHL	106,655	45,638	39,765	10,764	10,488
Australia	AUS	91,689	27,888	30,306	15,871	17,624
Mexico	MEX	79,033	18,441	29,434	12,256	18,902
Brazil	BRA	78,175	16,479	18,923	17,399	25,374
Italy	ITA	72,096	26,706	25,112	9,593	10,685
Germany	DEU	69,791	12,312	20,439	13,802	23,238
France	FRA	68,442	20,100	22,639	11,997	13,706
Colombia	COL	47,160	15,235	16,780	6,814	8,331
Spain	ESP	45,548	9,911	11,776	10,414	13,447
Taiwan	TWN	34,100	5,419	6,954	9,623	12,104
Japan	JPN	33,043	5,253	9,192	7,822	10,776
Korea, Rep.	KOR	28,889	6,035	8,437	6,033	8,384
Russian Federation	RUS	28,422	8,658	8,825	6,114	4,825
Puerto Rico	PRI	27,543	9,225	10,281	3,714	4,323
Pakistan	PAK	25,985	3,991	11,429	2,786	7,779
Egypt, Arab Rep.	EGY	25,061	3,935	10,035	3,402	7,689
Saudi Arabia	SAU	25,051	1,050	2,033	6,274	15,694
Singapore	SGP	24,869	3,302	4,318	7,691	9,558
Ireland	IRL	24,692	7,371	10,582	3,150	3,589
Argentina	ARG	23,984	6,722	8,119	3,574	5,569
Iran	IRN	21,379	4,426	7,915	3,176	5,862
Israel	ISR	20,642	5,743	8,576	2,564	3,759
Nigeria	NGA	20,541	4,629	7,534	3,157	5,221
South Africa	ZAF	20,168	5,547	6,894	3,568	4,159
Georgia	GEO	19,129	5,588	4,607	4,631	4,303
Venezuela, RB	VEN	18,959	6,195	6,973	2,599	3,192
Switzerland	CHE	18,821	3,378	4,403	4,963	6,077
Netherlands	NLD	18,500	4,242	6,012	3,536	4,710
Peru	PER	17,895	4,513	4,519	3,704	5,159
Hong Kong SAR, China	HKG	16,453	2,852	3,055	5,156	5,390
Turkiye	TUR	15,731	2,330	5,035	2,710	5,656
Bangladesh	BGD	15,344	3,792	5,509	2,197	3,846
New Zealand	NZL	15,083	4,307	5,335	2,485	2,956
Greece	GRC	14,915	2,973	4,201	3,280	4,461
Jordan	JOR	14,914	4,697	5,132	1,694	3,391
Poland	POL	14,367	5,465	3,992	2,629	2,281
Sweden	SWE	14,353	2,906	3,854	3,389	4,204
Ukraine	UKR	14,095	5,286	5,962	1,522	1,325
Lebanon	LBN	13,434	3,421	4,726	1,961	3,326
Chile	CHL	11,830	1,753	3,188	2,125	4,764
Belgium	BEL	11,760	3,054	3,993	1,914	2,799
Thailand	THA	10,969	4,312	3,314	1,827	1,516
Costa Rica	CRI	10,940	2,852	3,800	1,759	2,529
Dominican Republic	DOM	9,589	2,905	2,861	1,833	1,990
Denmark	DNK	9,464	1,818	2,440	2,056	3,150
Viet Nam	VNM	9,111	1,272	1,444	3,226	3,169
Malaysia	MYS	8,785	1,272	1,828	2,391	3,294
Austria	AUT	8,693	1,537	1,971	2,066	3,119
Norway	NOR	7,763	1,153	1,284	2,637	2,689
United Arab Emirates	ARE	7,438	1,087	1,178	2,106	3,067
Ecuador	ECU	7,343	1,823	1,830	1,554	2,136
Kenya	KEN	7,152	1,986	2,213	1,296	1,657
Romania	ROU	7,041	2,902	2,153	1,127	859
Indonesia	IDN	6,786	1,375	1,169	2,013	2,229
Portugal	PRT	6,761	1,454	1,901	1,422	1,984
Hungary	HUN	5,830	2,321	2,120	697	692
Ghana	GHA	5,598	1,056	1,990	907	1,645
Syrian Arab Republic	SYR	4,913	1,106	2,484	390	933
Jamaica	JAM	4,783	1,354	937	1,387	1,105
Trinidad and Tobago	TTO	4,761	751	658	1,734	1,618
Iraq	IRQ	4,507	679	2,151	357	1,320
Belarus	BLR	4,256	1,733	1,604	640	279
Bulgaria	BGR	4,159	1,480	1,336	753	590

Continued on next page

DRAFT

Guatemala	GTM	4,123	802	1,107	826	1,388
Panama	PAN	4,122	653	568	1,345	1,556
Czechia	CZE	3,954	1,165	1,039	908	842
Finland	FIN	3,920	1,568	791	982	579
Cuba	CUB	3,632	1,191	1,835	261	345
Serbia	SRB	3,321	1,012	1,215	554	540
Nepal	NPL	3,205	469	820	717	1,199
Bolivia	BOL	3,014	530	590	850	1,044
Uganda	UGA	2,931	718	882	487	844
Honduras	HND	2,896	754	858	591	693
Sri Lanka	LKA	2,836	633	897	580	726
El Salvador	SLV	2,663	702	867	484	610
Ethiopia	ETH	2,450	501	1,011	339	599
Lithuania	LTU	2,216	1,077	605	353	181
Armenia	ARM	2,117	755	689	383	290
Slovak Republic	SVK	2,068	919	596	302	251
Morocco	MAR	2,015	295	507	542	671
Albania	ALB	1,783	696	492	332	263
Cameroon	CMR	1,710	463	631	228	388
Guam	GUM	1,661	431	367	413	450
Nicaragua	NIC	1,645	320	497	326	502
West Bank and Gaza	PSE	1,631	394	598	270	369
Afghanistan	AFG	1,620	156	741	112	611
Kuwait	KWT	1,582	173	298	316	795
Zimbabwe	ZWE	1,577	374	476	283	444
Sudan	SDN	1,558	278	613	179	488
Guyana	GUY	1,551	188	159	604	600
Uruguay	URY	1,532	303	415	338	476
Bahrain	BHR	1,510	178	185	580	567
Macao SAR, China	MAC	1,396	703	129	370	194
Bermuda	BMU	1,388	169	138	589	492
Kazakhstan	KAZ	1,368	393	353	385	237
Moldova	MDA	1,251	600	326	205	120
Croatia	HRV	1,171	389	332	200	250
Azerbaijan	AZE	1,082	217	396	167	302
Latvia	LVA	1,066	495	251	218	102
Virgin Islands (U.S.)	VIR	1,042	305	185	348	204
Qatar	QAT	1,031	147	190	245	449
Cyprus	CYP	1,013	132	280	223	378
Paraguay	PRY	986	223	232	220	311
Algeria	DZA	968	208	381	109	270
Liberia	LBR	961	110	265	235	351
Tanzania	TZA	933	133	179	266	355
Bahamas, The	BHS	930	110	125	353	342
Niger	NER	929	338	260	189	142
Tunisia	TUN	918	124	218	251	325
Iceland	ISL	902	192	203	259	248
Oman	OMN	898	94	137	226	441
Bosnia and Herzegovina	BIH	867	217	225	220	205
Barbados	BRB	863	270	254	176	163
Haiti	HTI	772	134	160	205	273
Estonia	EST	693	264	188	128	113
Senegal	SEN	606	134	245	97	130
Malta	MLT	572	148	172	114	138
Belize	BLZ	557	91	60	213	193
Rwanda	RWA	546	84	87	163	212
Slovenia	SVN	518	178	127	110	103
Cambodia	KHM	510	84	99	171	156
Libya	LBY	494	43	140	60	251
North Macedonia	MKD	493	172	208	47	66
Uzbekistan	UZB	474	98	195	73	108
Yemen, Rep.	YEM	468	48	168	69	183
Myanmar	MMR	467	52	105	85	225
Mauritius	MUS	405	134	144	55	72
Zambia	ZMB	396	58	99	66	173
Luxembourg	LUX	382	63	65	129	125
Kyrgyz Republic	KGZ	329	104	80	98	47
Cayman Islands	CYM	321	35	26	160	100
Angola	AGO	287	25	44	71	147
Malawi	MWI	281	38	79	52	112
Montenegro	MNE	280	114	93	47	26
Namibia	NAM	270	41	40	94	95
Fiji	FJI	259	82	72	52	53

Continued on next page

DRAFT

Antigua and Barbuda	ATG	226	40	40	90	56
St. Lucia	LCA	213	24	30	102	57
Mongolia	MNG	207	39	38	80	50
Sierra Leone	SLE	185	23	66	29	67
American Samoa	ASM	185	20	25	85	55
St. Kitts and Nevis	KNA	184	45	35	55	49
Mozambique	MOZ	154	14	24	40	76
Cote dIvoire	CIV	153	11	27	48	67
Eritrea	ERI	150	15	47	22	66
Dominica	DMA	144	14	12	73	45
Monaco	MCO	141	22	26	51	42
Benin	BEN	139	26	56	10	47
Botswana	BWA	129	21	27	41	40
British Virgin Islands	VGB	120	15	10	58	37
Gambia, The	GMB	114	14	31	18	51
Madagascar	MDG	112	29	24	36	23
Papua New Guinea	PNG	96	32	15	21	28
Grenada	GRD	92	10	21	27	34
Northern Mariana Islands	MNP	91	8	3	51	29
Burkina Faso	BFA	90	10	25	26	29
Micronesia, Fed. Sts.	FSM	89	3	9	35	42
Somalia	SOM	86	6	27	7	46
Aruba	ABW	82	12	16	25	29
Brunei Darussalam	BRN	82	16	16	31	19
Gabon	GAB	80	17	16	23	24
Congo, Rep.	COG	74	7	19	18	30
Turks and Caicos Islands	TCA	73	9	7	29	28
Tajikistan	TJK	73	18	20	12	23
Sint Maarten (Dutch part)	SXM	72	6	8	35	23
San Marino	SMR	71	11	7	24	29
Congo, Dem. Rep.	COD	63	2	15	15	31
Togo	TGO	60	10	11	12	27
Mali	MLI	58	9	13	15	21
South Sudan	SSD	55	11	28	3	13
Cabo Verde	CPV	53	7	12	13	21
Suriname	SUR	51	7	6	16	22
Liechtenstein	LIE	49	9	16	13	11
Maldives	MDV	47	8	10	5	24
Marshall Islands	MHL	46	3	7	20	16
Guinea	GIN	43	5	9	12	17
Andorra	AND	42	9	12	10	11
Laos	LAO	42	2	7	12	21
St. Vincent and the Grenadines	VCT	40	5	17	13	5
New Caledonia	NCL	37	11	9	6	11
Lesotho	LSO	36	7	4	11	14
Turkmenistan	TKM	36	7	2	19	8
Djibouti	DJI	35	5	8	6	16
Curacao	CUW	34	4	5	9	16
Burundi	BDI	33	4	6	10	13
Gibraltar	GIB	31	6	9	5	11
Equatorial Guinea	GNQ	27	4	3	5	15
Central African Republic	CAF	25	4	14	0	7
Seychelles	SYC	23	4	3	6	10
Samoa	WSM	23	0	7	7	9
French Polynesia	PYF	22	9	4	3	6
Mauritania	MRT	20	2	2	6	10
Palau	PLW	18	0	3	7	8
Chad	TCD	17	0	3	6	8
Tonga	TON	16	2	1	7	6
Greenland	GRL	15	1	7	2	5
Solomon Islands	SLB	12	1	2	2	7
Bhutan	BTN	10	1	1	5	3
Vanuatu	VUT	10	0	3	4	3
Comoros	COM	8	0	1	2	5
Faroe Islands	FRO	8	3	1	4	0
Guinea-Bissau	GNB	8	1	1	3	3
St. Martin (French part)	MAF	7	3	0	2	2
Korea, Dem. People's Rep.	PRK	3	0	1	0	2
Tuvalu	TUV	1	0	0	1	0
Nauru	NRU	1	1	0	0	0
Kiribati	KIR	1	0	1	0	0

DRAFT



Fig. S2. Return, stay, and onward migration by gender in the twenty largest origin countries. Stacked bars show, for men and women separately in each origin country, the share of U.S. skilled immigrants whose last observed status is staying in the United States, returning to the country of origin, or moving to a third country (Revelio resume data, 1990–2019).

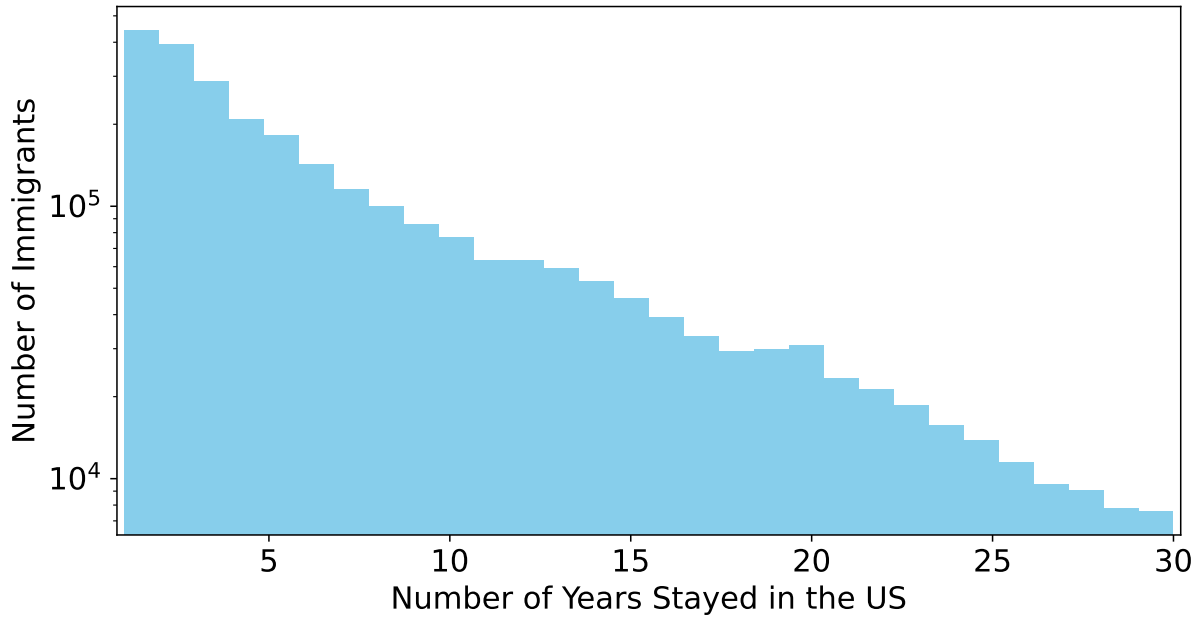


Fig. S3. Distribution of duration of U.S. stay among skilled immigrants. Histogram of the number of years each immigrant is observed in the United States; the x-axis reports years spent in the U.S., and the y-axis reports the number of immigrants on a logarithmic scale.

DRAFT

Table S3. Immigrants by last U.S. state of residence and gender. Counts of skilled immigrants by last observed U.S. state, disaggregated by gender (female, male) and shown alongside totals. "Unknown" denotes cases in which the state of the last U.S. record could not be determined.

Last U.S. State	Female	Male	Total	Last U.S. State	Female	Male	Total
Unknown	209,900	305,646	515,546	AL	5,936	7,642	13,578
CA	130,902	177,973	308,875	SC	5,215	6,746	11,961
NY	135,233	150,302	285,535	UT	4,520	6,849	11,369
IL	86,577	108,233	194,810	NM	4,057	5,752	9,809
TX	57,635	93,734	151,369	NV	4,234	5,294	9,528
FL	54,150	62,614	116,764	WV	3,143	5,019	8,162
MA	45,198	61,289	106,487	OK	3,376	4,622	7,998
PA	37,447	47,145	84,592	IA	3,298	4,557	7,855
WA	30,664	42,351	73,015	VT	3,948	3,766	7,714
NJ	31,113	39,824	70,937	KS	2,817	4,264	7,081
VA	26,248	31,281	57,529	DE	3,078	3,876	6,954
MI	18,580	30,780	49,360	AR	3,009	3,897	6,906
GA	21,110	27,055	48,165	NH	3,148	3,492	6,640
NC	19,858	26,279	46,137	RI	2,699	3,231	5,930
MD	19,511	21,988	41,499	HI	2,902	2,783	5,685
OH	16,993	24,335	41,328	NE	2,114	3,015	5,129
MO	16,517	18,921	35,438	ME	2,440	2,537	4,977
AZ	13,978	17,163	31,141	MS	1,919	2,354	4,273
CO	12,710	17,507	30,217	ID	1,592	2,327	3,919
IN	9,872	15,073	24,945	ND	1,562	2,085	3,647
CT	9,885	11,762	21,647	MT	1,537	1,611	3,148
OR	9,400	11,907	21,307	AK	1,505	1,620	3,125
DC	9,497	10,249	19,746	WY	998	1,289	2,287
WI	8,250	11,094	19,344	SD	853	1,020	1,873
MN	8,951	9,668	18,619	SP	9	18	27
KY	6,652	9,983	16,635	MP	9	3	12
TN	7,119	9,466	16,585	PR	5	1	6
LA	6,143	7,557	13,700				

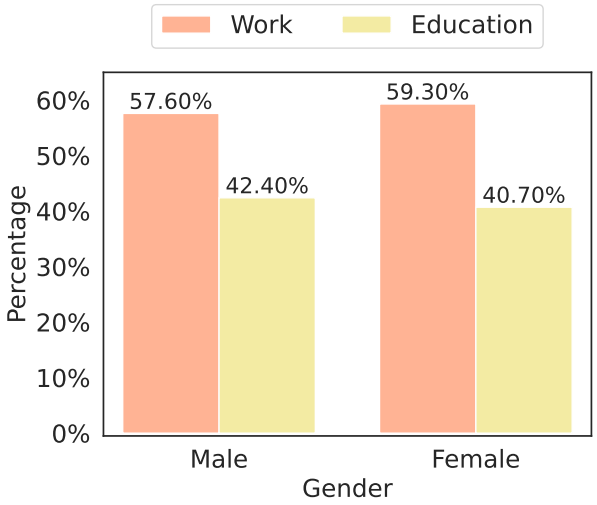


Fig. S4. Entry route into the United States by gender. Share of skilled immigrants whose first recorded experience in the U.S. is work versus education, shown separately for men and women.

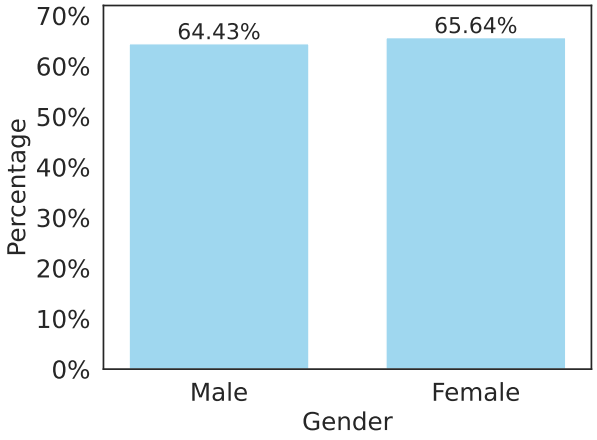


Fig. S5. U.S. work experience prior to last outcome, by gender (both cohorts). Percentage of skilled immigrants who held at least one job in the United States before their last recorded status (Stayed, Returned, or Moved), shown separately for men and women. Percentages are displayed above each bar.

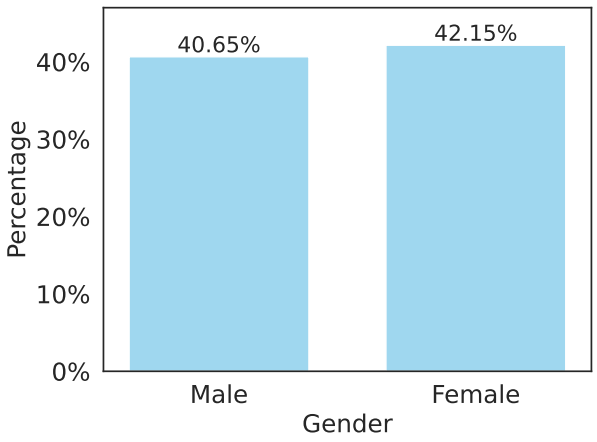


Fig. S6. U.S. work experience prior to last outcome, by gender (student cohort only). Percentage of skilled immigrants who held at least one job in the United States before their last recorded status (Stayed, Returned, or Moved), shown separately for men and women. Percentages are displayed above each bar.

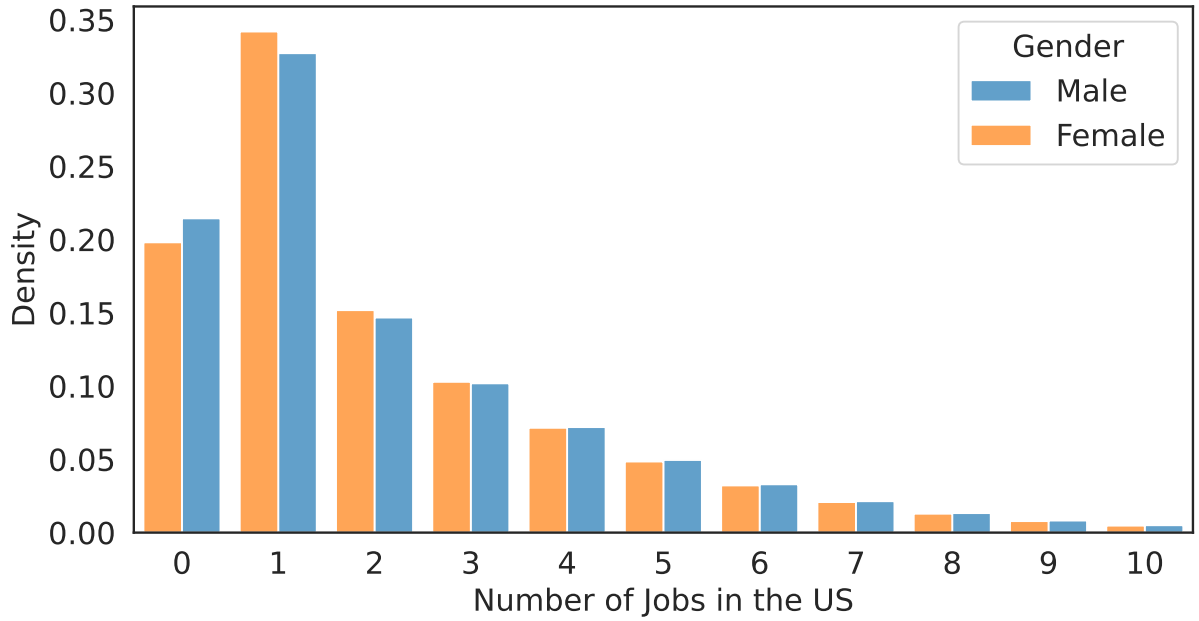


Fig. S7. Distribution of number of U.S. jobs by gender. Bars show the share of immigrants with 0, 1, 2, ... job records in the United States, separately for men and women. For visual clarity, the x-axis is truncated at 10 jobs; a small number of individuals hold up to 20 jobs.

Table S4. Highest degree attainment among skilled immigrants by gender. Entries report the share of female and male immigrants whose highest recorded qualification is a bachelor's, master's, MBA, or doctoral degree.

Degree	Female	Male
Bachelor	45.21%	41.22%
Master	35.55%	33.52%
MBA	9.16%	12.97%
Doctor	10.08%	12.30%

Table S5. Author counts by country of origin, immigrant status, gender, and immigrant–non-immigrant ratio. For each country of origin, the table reports the number of linked immigrant authors (from the SciSciNet–Revelio match) and non-immigrant authors (whose careers are entirely within a single country), disaggregated by gender. The final column gives the ratio of total immigrant to total non-immigrant authors for each origin country.

Country of Origin		Immigrants			Non-Immigrants			Immigrants÷Non
Name	ISO3C	Female	Male	Total	Female	Male	Total	Total÷Total
Afghanistan	AFG	5	29	34	184	599	783	0.043
Albania	ALB	28	20	48	1340	1770	3110	0.015
Algeria	DZA	6	20	26	5190	14771	19961	0.001
Angola	AGO	2	2	4	258	1110	1368	0.003
Argentina	ARG	582	968	1550	51478	70751	122229	0.013
Armenia	ARM	98	108	206	1191	7174	8365	0.025
Australia	AUS	2480	3093	5573	180749	339853	520602	0.011
Austria	AUT	242	516	758	38887	95663	134550	0.006
Azerbaijan	AZE	6	29	35	941	7438	8379	0.004
Bahrain	BHR	20	30	50	5632	12989	18621	0.003
Bangladesh	BGD	572	1232	1804	16748	63026	79774	0.023
Barbados	BRB	18	17	35	15	28	43	0.814
Belarus	BLR	61	51	112	1847	18652	20499	0.005
Belgium	BEL	369	562	931	53116	125312	178428	0.005
Benin	BEN	2	4	6	1084	4503	5587	0.001
Bhutan	BTN	0	0	0	94	164	258	0.000
Bolivia	BOL	41	60	101	591	1069	1660	0.061
Bosnia and Herzegovina	BIH	18	18	36	3684	3960	7644	0.005
Botswana	BWA	2	1	3	566	2572	3138	0.001
Brazil	BRA	1993	3086	5079	484726	592127	1076853	0.005
Bulgaria	BGR	76	81	157	10459	26793	37252	0.004
Burkina Faso	BFA	1	5	6	194	1114	1308	0.005
Burundi	BDI	0	0	0	29	224	253	0.000

Continued on next page

DRAFT

Cabo Verde	CPV	0	2	2	43	76	119	0.017
Cambodia	KHM	9	9	18	3088	6731	9819	0.002
Cameroon	CMR	14	38	52	3070	13835	16905	0.003
Canada	CAN	6858	8821	15679	281668	581866	863534	0.018
Central African Republic	CAF	0	1	1	0	0	0	NaN
Chad	TCD	0	0	0	6	153	159	0.000
Chile	CHL	329	692	1021	36935	73050	109985	0.009
China	CHN	10000	12425	22425	2374787	5180641	7555428	0.003
Colombia	COL	1167	1838	3005	56274	97115	153389	0.020
Comoros	COM	1	0	1	0	0	0	NaN
Congo, Dem. Rep.	COD	0	0	0	472	3262	3734	0.000
Congo, Rep.	COG	3	1	4	102	781	883	0.005
Costa Rica	CRI	305	556	861	562	1178	1740	0.495
Cote d'Ivoire	CIV	2	1	3	570	3279	3849	0.001
Croatia	HRV	52	53	105	13897	14593	28490	0.004
Cuba	CUB	86	122	208	3647	6568	10215	0.020
Cyprus	CYP	34	46	80	2241	3919	6160	0.013
Czechia	CZE	109	159	268	30826	74940	105766	0.003
Denmark	DNK	245	476	721	45572	85372	130944	0.006
Dominican Republic	DOM	79	109	188	773	1061	1834	0.103
Ecuador	ECU	135	219	354	11128	24633	35761	0.010
Egypt, Arab Rep.	EGY	618	2082	2700	55643	126922	182565	0.015
El Salvador	SLV	22	33	55	1367	2119	3486	0.016
Eritrea	ERI	0	0	0	5	80	85	0.000
Estonia	EST	21	36	57	6036	11661	17697	0.003
Eswatini	SWZ	0	0	0	102	398	500	0.000
Ethiopia	ETH	39	127	166	5875	25789	31664	0.005
Fiji	FJI	10	12	22	384	1168	1552	0.014
Finland	FIN	174	99	273	39819	70101	109920	0.002
France	FRA	977	1796	2773	235373	664043	899416	0.003
Gabon	GAB	0	3	3	196	587	783	0.004
Gambia, The	GMB	2	3	5	29	129	158	0.032
Georgia	GEO	291	309	600	1352	3928	5280	0.114
Germany	DEU	1733	4125	5858	337153	1004652	1341805	0.004
Ghana	GHA	94	307	401	6826	21877	28703	0.014
Greece	GRC	595	1224	1819	31352	87411	118763	0.015
Guatemala	GTM	39	60	99	932	1498	2430	0.041
Guinea	GIN	1	0	1	3	16	19	0.053
Guinea-Bissau	GNB	0	0	0	36	72	108	0.000
Guyana	GUY	16	8	24	108	169	277	0.087
Haiti	HTI	7	8	15	7	34	41	0.366
Honduras	HND	39	56	95	669	1111	1780	0.053
Hong Kong SAR, China	HKG	653	707	1360	38342	113136	151478	0.009
Hungary	HUN	185	274	459	20407	45090	65497	0.007
Iceland	ISL	41	46	87	1880	3012	4892	0.018
India	IND	9350	19902	29252	273753	1060374	1334127	0.022
Indonesia	IDN	144	231	375	132816	243413	376229	0.001
Iran	IRN	1781	3891	5672	166982	364711	531693	0.011
Iraq	IRQ	46	272	318	9883	21364	31247	0.010
Ireland	IRL	596	957	1553	31418	68648	100066	0.016
Israel	ISR	756	1223	1979	50620	99397	150017	0.013
Italy	ITA	3955	4486	8441	192918	385976	578894	0.015
Jamaica	JAM	51	71	122	2767	5704	8471	0.014
Japan	JPN	857	1903	2760	316220	1276386	1592606	0.002
Jordan	JOR	285	725	1010	7365	18044	25409	0.040
Kazakhstan	KAZ	32	32	64	3263	9523	12786	0.005
Kenya	KEN	109	222	331	6480	18842	25322	0.013
Korea, Rep.	KOR	1278	1998	3276	163628	508846	672474	0.005
Kuwait	KWT	49	85	134	2900	10637	13537	0.010
Kyrgyz Republic	KGZ	3	5	8	311	1139	1450	0.006
Lao PDR	LAO	0	1	1	230	981	1211	0.001
Latvia	LVA	22	11	33	5717	9558	15275	0.002
Lebanon	LBN	490	765	1255	6831	11090	17921	0.070
Lesotho	LSO	2	1	3	72	351	423	0.007
Liberia	LBR	4	9	13	17	83	100	0.130
Libya	LBY	15	66	81	860	2107	2967	0.027
Lithuania	LTU	40	31	71	12600	14016	26616	0.003
Luxembourg	LUX	7	8	15	1367	4039	5406	0.003
Madagascar	MDG	3	2	5	348	1103	1451	0.003
Malawi	MWI	3	16	19	576	1764	2340	0.008
Malaysia	MYS	155	258	413	58110	159534	217644	0.002
Maldives	MDV	0	1	1	33	56	89	0.011
Mali	MLI	0	0	0	232	2093	2325	0.000

Continued on next page

DRAFT

Malta	MLT	16	39	55	1143	2024	3167	0.017
Mauritania	MRT	0	0	0	14	172	186	0.000
Mauritius	MUS	7	5	12	448	1489	1937	0.006
Mexico	MEX	736	1538	2274	96785	215997	312782	0.007
Moldova	MDA	14	8	22	498	2282	2780	0.008
Mongolia	MNG	2	7	9	555	2900	3455	0.003
Montenegro	MNE	4	6	10	611	1035	1646	0.006
Morocco	MAR	37	26	63	3989	11729	15718	0.004
Mozambique	MOZ	3	5	8	6279	13851	20130	0.000
Myanmar	MMR	7	15	22	829	1142	1971	0.011
Namibia	NAM	3	4	7	0	0	0	NaN
Nepal	NPL	126	332	458	3807	11521	15328	0.030
Netherlands	NLD	519	882	1401	98395	285667	384062	0.004
New Zealand	NZL	422	559	981	27483	59648	87131	0.011
Nicaragua	NIC	9	39	48	1225	2934	4159	0.012
Niger	NER	12	17	29	129	803	932	0.031
Nigeria	NGA	365	819	1184	21760	111427	133187	0.009
North Macedonia	MKD	2	5	7	1699	1850	3549	0.002
Norway	NOR	181	216	397	33083	65426	98509	0.004
Oman	OMN	24	43	67	2684	6415	9099	0.007
Pakistan	PAK	497	1788	2285	47455	104369	151824	0.015
Panama	PAN	46	84	130	863	1599	2462	0.053
Papua New Guinea	PNG	1	2	3	310	1080	1390	0.002
Paraguay	PRY	9	24	33	1219	1472	2691	0.012
Peru	PER	238	463	701	13496	26482	39978	0.018
Philippines	PHL	1838	2296	4134	16360	30807	47167	0.088
Poland	POL	461	471	932	88513	150348	238861	0.004
Portugal	PRT	322	487	809	40308	61469	101777	0.008
Qatar	QAT	25	35	60	3401	7907	11308	0.005
Romania	ROU	232	224	456	23240	40377	63617	0.007
Russian Federation	RUS	829	1411	2240	86416	571163	657579	0.003
Rwanda	RWA	10	15	25	284	1341	1625	0.015
Saudi Arabia	SAU	515	1283	1798	22851	57957	80808	0.022
Senegal	SEN	4	6	10	654	3121	3775	0.003
Serbia	SRB	129	193	322	14131	25237	39368	0.008
Seychelles	SYC	1	2	3	9	16	25	0.120
Sierra Leone	SLE	1	5	6	77	397	474	0.013
Singapore	SGP	850	1337	2187	37943	111737	149680	0.015
Slovak Republic	SVK	39	49	88	10589	20758	31347	0.003
Slovenia	SVN	21	22	43	8331	15139	23470	0.002
Somalia	SOM	0	1	1	13	85	98	0.010
South Africa	ZAF	461	743	1204	39287	111625	150912	0.008
South Sudan	SSD	0	1	1	0	0	0	NaN
Spain	ESP	1617	2125	3742	189807	399891	589698	0.006
Sri Lanka	LKA	134	135	269	3599	24807	28406	0.009
Sudan	SDN	30	126	156	2138	5981	8119	0.019
Suriname	SUR	2	1	3	90	202	292	0.010
Sweden	SWE	403	757	1160	68951	134342	203293	0.006
Switzerland	CHE	425	870	1295	69020	185120	254140	0.005
Syrian Arab Republic	SYR	29	181	210	893	3247	4140	0.051
Tajikistan	TJK	0	0	0	84	1390	1474	0.000
Tanzania	TZA	22	39	61	1861	6663	8524	0.007
Thailand	THA	127	108	235	29464	92021	121485	0.002
Timor-Leste	TLS	0	0	0	2	9	11	0.000
Togo	TGO	1	1	2	113	1304	1417	0.001
Trinidad and Tobago	TTO	67	71	138	151	474	625	0.221
Tunisia	TUN	15	35	50	12990	17734	30724	0.002
Turkiye	TUR	482	1315	1797	96748	177833	274581	0.007
Turkmenistan	TKM	1	1	2	1	71	72	0.028
Uganda	UGA	65	127	192	3518	8650	12168	0.016
Ukraine	UKR	167	197	364	26580	106093	132673	0.003
United Arab Emirates	ARE	150	210	360	7263	16202	23465	0.015
United Kingdom	GBR	5007	8471	13478	451553	1234978	1686531	0.008
Uruguay	URY	40	64	104	4979	6515	11494	0.009
Uzbekistan	UZB	4	8	12	688	8107	8795	0.001
Vanuatu	VUT	0	0	0	2	4	6	0.000
Venezuela, RB	VEN	242	424	666	6449	13514	19963	0.033
Viet Nam	VNM	347	452	799	13293	32535	45828	0.017
West Bank and Gaza	PSE	22	60	82	0	1	1	82.000
Yemen, Rep.	YEM	6	17	23	440	1949	2389	0.010
Zambia	ZMB	2	18	20	1450	2556	4006	0.005
Zimbabwe	ZWE	37	82	119	2040	9203	11243	0.011

DRAFT

S3. Representativeness Analyses

Publicly available data on skilled immigrants to the United States are fragmented across administrative systems and are not designed to follow individuals' careers. To assess how well the Revelio resume data capture the broader population of skilled immigrants, we therefore compare our sample to two external benchmarks and treat them as envelopes within which our resume-based estimates should lie.

First, we use annual counts of F–1 student visas and H–1B specialty-occupation visas issued by the U.S. Department of State between 1997 and 2019.[¶¶] These "nonimmigrant visa statistics" are the main public series on high-skilled temporary migration, but they record visa issuances and petition approvals rather than unique individuals. Many H–1B approvals are granted to workers who are already in the United States because they extend their stay, change employers, or change status from another visa (for example, from F–1 to H–1B). Likewise, a single person can receive multiple F–1 visas over time if they change degree programs or institutions. As a result, the official F–1 and H–1B counts are best interpreted as an *upper bound* on the number of people who could appear in a given arrival cohort, and our coverage ratios using these series are deliberately **conservative**: they tend to understate the share of skilled immigrants observed in the resume data.

Second, we benchmark our student subsample against the Student and Exchange Visitor Information System (SEVIS), which records "active" F–1 and M–1 student records at SEVP-certified schools in the United States.[***] SEVIS snapshots count all students who are in active status in a given month, including continuing students who arrived in earlier years, students on post-completion Optional Practical Training (OPT), and in some cases multiple records associated with the same person as they change programs or institutions. These data therefore measure a stock of active student records, not the flow of new international students entering in that year. Consequently, SEVIS totals will generally exceed the number of distinct individuals in any single immigration cohort, and our SEVIS-based coverage estimates should likewise be interpreted as lower bounds on the true fraction of international students represented in the resume sample.

Despite these limitations, the visa and SEVIS statistics provide two crucial benchmarks. The visa series allow us to show that the overall size of our immigrant sample tracks the growth and contraction of high-skilled inflows over time, while SEVIS is the only public source that offers gender-disaggregated counts for international students in recent years. By construction, neither source contains detailed information on education histories, employers, job titles, or subsequent migration decisions, and gender breakdowns are unavailable for H–1B workers. The Revelio resume data therefore play a complementary role: they sit inside the broad envelopes defined by visa and SEVIS counts but add the individual-level trajectories and gender information that are missing from existing public datasets.

Figures S8–S10 and Table S6 summarize these representativeness checks. Figure S8 reports yearly coverage rates by visa category, comparing the share of all H–1B+F–1 visa issuances, H–1B worker visas, and F–1 student visas that we observe in the corresponding resume subsamples. Figure S9 shows the underlying counts, documenting that the time series of immigrants in the resume data scales with the population of high-skilled visa issuances while remaining below them in level, consistent with the administrative over-counting of individuals. Figure S10 focuses on students and uses SEVIS to compute gender-specific coverage rates, providing the only available benchmark for the gender composition of our student subsample. Finally, Table S6 assesses cross-sectional representativeness by country of origin: for each country, we compute its share of all H–1B+F–1 visas and its share of immigrants in the Revelio sample over the same period, and summarize their ratio as a country-level representation measure. Values of this ratio below 100% indicate that immigrants from that country are underrepresented in the resume data relative to their share in the visa population, while values above 100% indicate overrepresentation. Together, Figures S8–S10 and Table S6 show that coverage improves over time, varies by visa category and gender, and is generally higher for English-speaking and high-income sending countries than for large sending countries such as China and India.

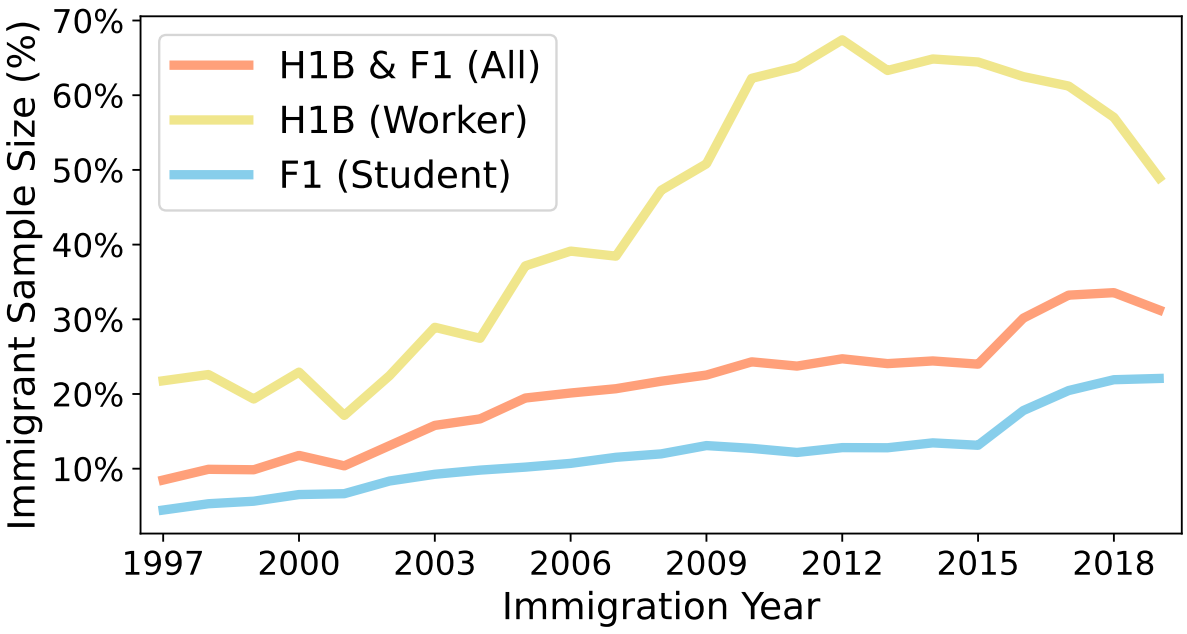


Fig. S8. Yearly representativeness. Percentage of U.S. skilled immigrants captured in the Revelio resume dataset by immigration year and visa type. "H1B & F1 (All)" compares the full resume sample to the combined number of H-1B and F-1 visas issued; "H1B (Worker)" and "F1 (Student)" restrict the comparison to worker and student subsamples, respectively. Government counts are based on non-immigrant visa statistics from the U.S. Department of State (1997–2019).

[¶¶] https://travel.state.gov/content/travel/en/legal/visa-law0/visa-statistics/nonimmigrant-visa-statistics.html

[***] SEVIS data are taken from the Department of Homeland Security's *SEVIS by the Numbers* reports for 2014–2019, available at https://www.ice.gov/sevis.

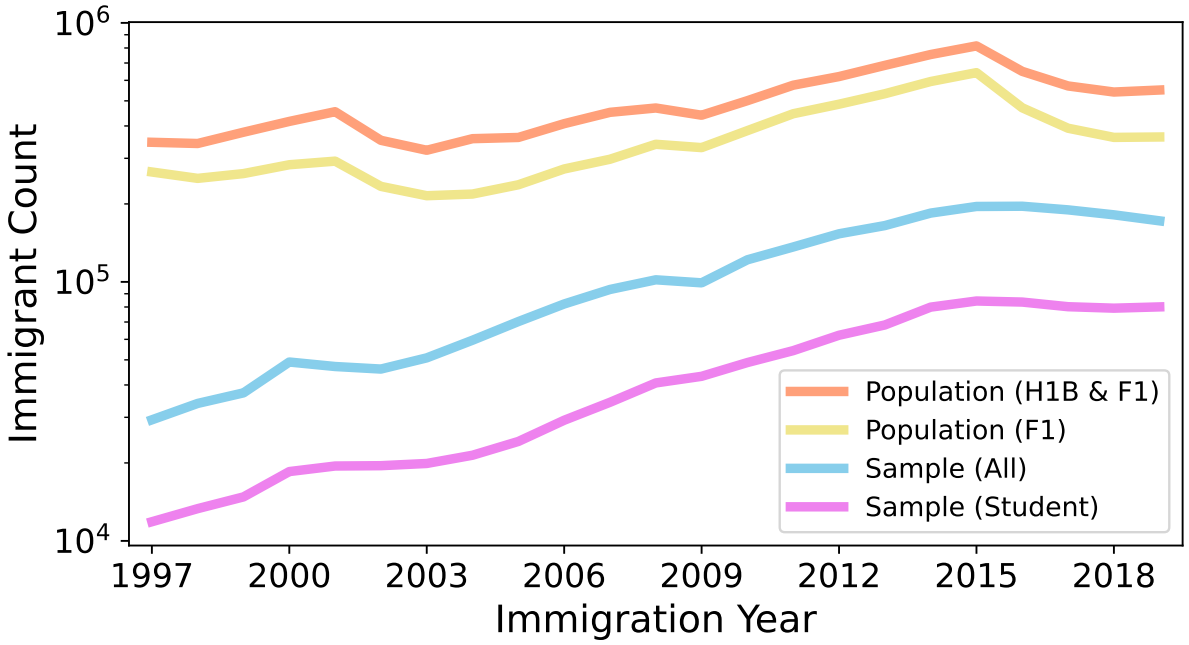


Fig. S9. Yearly coverage. Comparison of absolute counts: total H-1B+F-1 visa issuances, F-1 visa issuances alone, all immigrants observed in the Revelio resume data, and the student subsample, by immigration year. The figure shows that the resume sample scales with the underlying inflows of skilled immigrants over time, while remaining smaller than administrative totals because the latter include multiple issuances per person, dependents, and other records that are not present in resume data.

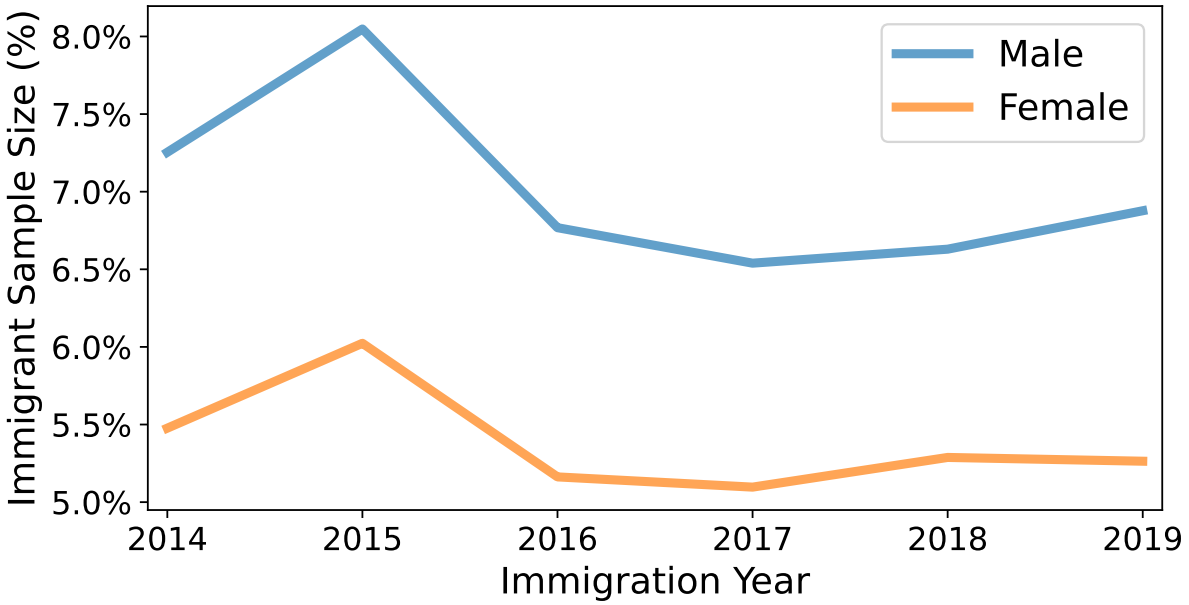


Fig. S10. Yearly representativeness by gender (students). Percentage of female and male F-1 students in the Revelio resume dataset relative to the number of active F-1 SEVIS records for each year from 2014 to 2019. SEVIS counts represent active student records (including continuing students and those on post-completion training), so the resulting sample shares should be interpreted as conservative lower bounds on gender-specific coverage.

Table S6. Country-level representativeness relative to H-1B and F-1 visa shares. Visa columns provide the total number and percentage of the H1B and F1 visas issued reported by the US Department of State between 1997 and 2019. The sample columns specify the number and percentage of the U.S. skilled immigrants from the resume dataset between 1997 and 2019. The Representation is obtained from Sample% divided by Visa%. A value below 100 means that the sample of the corresponding country is underrepresented relative to the visa percentage.

Country of Origin		Visa (H1B+F1)		Sample (Resume)		Representation
Name	ISO3C	Number	Percentage	Number	Percentage	Sample%/Visa%
India	IND	2,516,992	22.15%	379,070	15.20%	68.61%
China	CHN	2,305,903	20.30%	204,321	8.19%	40.37%
Korea, Rep.	KOR	805,567	7.09%	26,757	1.07%	15.13%
Japan	JPN	573,154	5.04%	30,241	1.21%	24.04%
Taiwan	TWN	330,615	2.91%	31,445	1.26%	43.33%
Brazil	BRA	296,895	2.61%	75,753	3.04%	116.24%
Saudi Arabia	SAU	278,410	2.45%	24,659	0.99%	40.35%
Mexico	MEX	257,009	2.26%	75,430	3.02%	133.70%
United Kingdom	GBR	216,131	1.90%	225,393	9.04%	475.09%
Germany	DEU	205,075	1.80%	65,228	2.62%	144.90%
Vietnam	VNM	181,827	1.60%	9,060	0.36%	22.70%
France	FRA	171,938	1.51%	65,432	2.62%	173.37%
Turkey	TUR	164,350	1.45%	14,762	0.59%	40.92%
Colombia	COL	130,260	1.15%	44,938	1.80%	157.16%
Russia	RUS	111,973	0.99%	26,151	1.05%	106.40%
Spain	ESP	106,049	0.93%	43,492	1.74%	186.83%
Thailand	THA	105,526	0.93%	10,195	0.41%	44.01%
Italy	ITA	104,500	0.92%	69,866	2.80%	304.58%
Venezuela	VEN	99,791	0.88%	17,921	0.72%	81.81%
Indonesia	IDN	91,846	0.81%	6,347	0.25%	31.48%
Switzerland	CHE	86,265	0.76%	17,937	0.72%	94.72%
Nigeria	NGA	81,635	0.72%	20,207	0.81%	112.77%
Philippines	PHL	80,677	0.71%	97,420	3.91%	550.11%
Sweden	SWE	62,998	0.55%	13,443	0.54%	97.21%
Pakistan	PAK	61,710	0.54%	24,846	1.00%	183.42%
Australia	AUS	61,116	0.54%	88,083	3.53%	656.58%
Malaysia	MYS	61,104	0.54%	8,515	0.34%	63.48%
Nepal	NPL	57,932	0.51%	3,178	0.13%	24.99%
Argentina	ARG	57,458	0.51%	22,252	0.89%	176.43%
Hong Kong SAR, China	HKG	52,191	0.46%	15,907	0.64%	138.85%
Israel	ISR	48,353	0.43%	19,426	0.78%	183.02%
Singapore	SGP	48,121	0.42%	24,247	0.97%	229.55%
Kenya	KEN	47,045	0.41%	6,990	0.28%	67.69%
Netherlands	NLD	42,798	0.38%	17,327	0.69%	184.44%
Kuwait	KWT	39,381	0.35%	1,472	0.06%	17.03%
Peru	PER	39,036	0.34%	17,148	0.69%	200.12%
Iran	IRN	37,152	0.33%	20,990	0.84%	257.38%
Norway	NOR	36,616	0.32%	7,384	0.30%	91.87%
Jamaica	JAM	35,421	0.31%	4,581	0.18%	58.92%
Chile	CHL	35,364	0.31%	11,365	0.46%	146.41%
Ecuador	ECU	33,636	0.30%	7,057	0.28%	95.58%
Bangladesh	BGD	33,611	0.30%	14,733	0.59%	199.69%
South Africa	ZAF	32,859	0.29%	18,979	0.76%	263.13%
Poland	POL	31,953	0.28%	13,695	0.55%	195.25%
Bulgaria	BGR	29,966	0.26%	4,014	0.16%	61.02%
Greece	GRC	28,764	0.25%	13,981	0.56%	221.43%
Bahamas, The	BHS	28,310	0.25%	900	0.04%	14.48%
Ireland	IRL	26,156	0.23%	23,202	0.93%	404.11%
Egypt	EGY	25,815	0.23%	23,989	0.96%	423.34%
Denmark	DNK	25,785	0.23%	9,042	0.36%	159.75%
Ukraine	UKR	25,189	0.22%	13,503	0.54%	244.21%
Kazakhstan	KAZ	24,770	0.22%	1,340	0.05%	24.64%
Ghana	GHA	24,312	0.21%	5,499	0.22%	103.04%
Jordan	JOR	24,292	0.21%	14,056	0.56%	263.60%
Romania	ROU	23,671	0.21%	6,679	0.27%	128.54%
Sri Lanka	LKA	23,492	0.21%	2,591	0.10%	50.25%
Trinidad and Tobago	TTO	22,718	0.20%	4,531	0.18%	90.86%
United Arab Emirates	ARE	22,718	0.20%	7,337	0.29%	147.13%
Belgium	BEL	21,671	0.19%	10,993	0.44%	231.09%
Austria	AUT	21,358	0.19%	8,246	0.33%	175.89%
Panama	PAN	19,960	0.18%	3,929	0.16%	89.67%
Lebanon	LBN	19,307	0.17%	13,007	0.52%	306.91%
Dominican Republic	DOM	18,210	0.16%	9,201	0.37%	230.18%
New Zealand	NZL	18,170	0.16%	14,264	0.57%	357.63%
Mongolia	MNG	18,139	0.16%	205	0.01%	5.15%
Morocco	MAR	17,241	0.15%	1,982	0.08%	52.37%

Continued on next page

DRAFT

Table S6. Country-level representativeness relative to H-1B and F-1 visa shares. Visa columns provide the total number and percentage of the H1B and F1 visas issued reported by the US Department of State between 1997 and 2019. The sample columns specify the number and percentage of the U.S. skilled immigrants from the resume dataset between 1997 and 2019. The Representation is obtained from Sample% divided by Visa%. A value below 100 means that the sample of the corresponding country is underrepresented relative to the visa percentage.

Country of Origin		Visa (H1B+F1)		Sample (Resume)		Representation
Name	ISO3C	Number	Percentage	Number	Percentage	Sample%/Visa%
Czech Republic	CZE	17,059	0.15%	3,807	0.15%	101.67%
Ethiopia	ETH	16,611	0.15%	2,408	0.10%	66.04%
Cameroon	CMR	15,562	0.14%	1,700	0.07%	49.77%
Portugal	PRT	14,990	0.13%	6,581	0.26%	200.00%
Zimbabwe	ZWE	14,150	0.12%	1,519	0.06%	48.90%
Hungary	HUN	14,009	0.12%	5,596	0.22%	181.98%
Serbia and Montenegro	SRB	13,968	0.12%	3,169	0.13%	103.36%
Honduras	HND	13,744	0.12%	2,773	0.11%	91.91%
Costa Rica	CRI	13,345	0.12%	10,219	0.41%	348.85%
Burma	MMR	13,295	0.12%	464	0.02%	15.90%
Tanzania	TZA	13,069	0.12%	899	0.04%	31.34%
Bolivia	BOL	12,354	0.11%	2,919	0.12%	107.64%
Finland	FIN	12,061	0.11%	3,736	0.15%	141.11%
Guatemala	GTM	11,775	0.10%	3,945	0.16%	152.63%
Cyprus	CYP	11,285	0.10%	920	0.04%	37.14%
Slovakia	SVK	11,204	0.10%	1,996	0.08%	81.16%
Angola	AGO	10,943	0.10%	278	0.01%	11.57%
Croatia	HRV	10,479	0.09%	1,111	0.04%	48.30%
Albania	ALB	10,441	0.09%	1,690	0.07%	73.74%
Uganda	UGA	10,268	0.09%	2,770	0.11%	122.90%
Qatar	QAT	10,245	0.09%	1,001	0.04%	44.51%
Oman	OMN	9,982	0.09%	880	0.04%	40.16%
El Salvador	SLV	9,228	0.08%	2,582	0.10%	127.47%
Libya	LBY	8,907	0.08%	486	0.02%	24.86%
Uzbekistan	UZB	7,845	0.07%	455	0.02%	26.42%
Haiti	HTI	7,807	0.07%	743	0.03%	43.36%
Senegal	SEN	7,796	0.07%	592	0.02%	34.59%
Belarus	BLR	7,775	0.07%	4,085	0.16%	239.35%
Syria	SYR	7,710	0.07%	4,766	0.19%	281.61%
Burkina Faso	BFA	7,541	0.07%	88	0.00%	5.32%
Azerbaijan	AZE	7,374	0.06%	1,051	0.04%	64.93%
Yemen	YEM	7,150	0.06%	455	0.02%	28.99%
Lithuania	LTU	7,029	0.06%	2,141	0.09%	138.76%
Georgia	GEO	7,017	0.06%	18,412	0.74%	1195.36%
Canada	CAN	6,713	0.06%	189,730	7.61%	12875.63%
Cambodia	KHM	6,686	0.06%	506	0.02%	34.48%
Rwanda	RWA	6,475	0.06%	531	0.02%	37.36%
Paraguay	PRY	6,323	0.06%	949	0.04%	68.37%
Iceland	ISL	5,911	0.05%	813	0.03%	62.66%
Cote d'Ivoire	CIV	5,726	0.05%	153	0.01%	12.17%
Barbados	BRB	5,708	0.05%	829	0.03%	66.16%
Mali	MLI	5,667	0.05%	54	0.00%	4.34%
Tunisia	TUN	5,665	0.05%	894	0.04%	71.89%
Uruguay	URY	5,473	0.05%	1,453	0.06%	120.95%
Bosnia and Herzegovina	BIH	5,145	0.05%	820	0.03%	72.61%
Iraq	IRQ	4,929	0.04%	4,423	0.18%	408.80%
Latvia	LVA	4,789	0.04%	972	0.04%	92.46%
Zambia	ZMB	4,771	0.04%	390	0.02%	37.24%
Tajikistan	TJK	4,719	0.04%	71	0.00%	6.85%
Congo, Dem. Rep. of the (Kinshasa)	COD	4,696	0.04%	63	0.00%	6.11%
Nicaragua	NIC	4,629	0.04%	1,593	0.06%	156.78%
Moldova	MDA	4,458	0.04%	1,212	0.05%	123.85%
Belize	BLZ	4,429	0.04%	527	0.02%	54.21%
Armenia	ARM	4,389	0.04%	2,046	0.08%	212.37%
Sudan	SDN	4,292	0.04%	1,504	0.06%	159.64%
Bahrain	BHR	4,202	0.04%	1,439	0.06%	156.01%
Gambia, The	GMB	3,975	0.03%	112	0.00%	12.84%
Kyrgyzstan	KGZ	3,926	0.03%	324	0.01%	37.60%
Slovenia	SVN	3,820	0.03%	484	0.02%	57.72%
Algeria	DZA	3,815	0.03%	954	0.04%	113.92%
Macedonia	MKD	3,661	0.03%	491	0.02%	61.10%
Togo	TGO	3,569	0.03%	59	0.00%	7.53%
Benin	BEN	3,323	0.03%	136	0.01%	18.64%
Dominica	DMA	3,207	0.03%	141	0.01%	20.03%
Gabon	GAB	3,166	0.03%	73	0.00%	10.50%
Niger	NER	3,117	0.03%	921	0.04%	134.61%
Turkmenistan	TKM	3,042	0.03%	36	0.00%	5.39%

Continued on next page

DRAFT

Table S6. Country-level representativeness relative to H-1B and F-1 visa shares. Visa columns provide the total number and percentage of the H1B and F1 visas issued reported by the US Department of State between 1997 and 2019. The sample columns specify the number and percentage of the U.S. skilled immigrants from the resume dataset between 1997 and 2019. The Representation is obtained from Sample% divided by Visa%. A value below 100 means that the sample of the corresponding country is underrepresented relative to the visa percentage.

Country of Origin		Visa (H1B+F1)		Sample (Resume)		Representation
Name	ISO3C	Number	Percentage	Number	Percentage	Sample%/Visa%
Guyana	GUY	3,016	0.03%	1,483	0.06%	224.01%
Afghanistan	AFG	2,985	0.03%	1,594	0.06%	243.27%
Liberia	LBR	2,981	0.03%	912	0.04%	139.37%
Estonia	EST	2,835	0.02%	670	0.03%	107.66%
Botswana	BWA	2,728	0.02%	122	0.00%	20.37%
Equatorial Guinea	GNQ	2,454	0.02%	27	0.00%	5.01%
Grenada	GRD	2,441	0.02%	90	0.00%	16.80%
Malawi	MWI	2,433	0.02%	267	0.01%	49.99%
Burundi	BDI	2,393	0.02%	33	0.00%	6.28%
Guinea	GIN	2,358	0.02%	42	0.00%	8.11%
Congo, Dem. Rep. of the (Congo Kinshasa)	COD	2,243	0.02%	63	0.00%	12.80%
Antigua and Barbuda	ATG	2,217	0.02%	222	0.01%	45.62%
Mauritius	MUS	2,100	0.02%	400	0.02%	86.77%
Eritrea	ERI	1,947	0.02%	150	0.01%	35.10%
Mauritania	MRT	1,850	0.02%	19	0.00%	4.68%
Madagascar	MDG	1,615	0.01%	109	0.00%	30.75%
Sierra Leone	SLE	1,579	0.01%	182	0.01%	52.51%
Bhutan	BTN	1,431	0.01%	9	0.00%	2.87%
Mozambique	MOZ	1,408	0.01%	147	0.01%	47.56%
Chad	TCD	1,404	0.01%	16	0.00%	5.19%
Laos	LAO	1,385	0.01%	41	0.00%	13.49%
Luxembourg	LUX	1,319	0.01%	375	0.02%	129.52%
Montenegro	MNE	1,211	0.01%	279	0.01%	104.96%
Congo, Rep. of the (Congo Brazzaville)	COG	1,153	0.01%	72	0.00%	28.45%
Fiji	FJI	1,093	0.01%	256	0.01%	106.70%
Congo, Rep. of the (Brazzaville)	COG	1,017	0.01%	72	0.00%	32.25%
Papua New Guinea	PNG	1,017	0.01%	88	0.00%	39.42%
Suriname	SUR	953	0.01%	50	0.00%	23.90%
Namibia	NAM	910	0.01%	259	0.01%	129.66%
Swaziland	SWZ	782	0.01%	0	-	-
Cuba	CUB	723	0.01%	3,516	0.14%	2215.44%
Tonga	TON	590	0.01%	15	0.00%	11.58%
Cape Verde	CPV	533	0.00%	52	0.00%	44.45%
Liechtenstein	LIE	493	0.00%	49	0.00%	45.28%
Samoa	WSM	470	0.00%	20	0.00%	19.39%
Central African Republic	CAF	440	0.00%	24	0.00%	24.85%
Malta	MLT	435	0.00%	550	0.02%	576.00%
Somalia	SOM	406	0.00%	78	0.00%	87.52%
South Sudan	SSD	369	0.00%	55	0.00%	67.90%
Lesotho	LSO	336	0.00%	29	0.00%	39.32%
Brunei	BRN	329	0.00%	81	0.00%	112.16%
Korea, North	PRK	296	0.00%	3	0.00%	4.62%
Maldives	MDV	210	0.00%	47	0.00%	101.96%
Andorra	AND	160	0.00%	42	0.00%	119.59%
Monaco	MCO	159	0.00%	132	0.01%	378.20%
Djibouti	DJI	150	0.00%	33	0.00%	100.22%
Kiribati	KIR	128	0.00%	1	0.00%	3.56%
Timor-Leste	TLS	124	0.00%	0	-	-
Guinea-Bissau	GNB	120	0.00%	8	0.00%	30.37%
Seychelles	SYC	90	0.00%	23	0.00%	116.42%
Solomon Islands	SLB	74	0.00%	11	0.00%	67.72%
San Marino	SMR	65	0.00%	68	0.00%	476.59%
Comoros	COM	56	0.00%	8	0.00%	65.08%
Sao Tome and Principe	STP	50	0.00%	0	-	-
Vanuatu	VUT	34	0.00%	10	0.00%	133.99%
Marshall Islands	MHL	31	0.00%	46	0.00%	676.00%
Nauru	NRU	12	0.00%	1	0.00%	37.96%
Tuvalu	TUV	9	0.00%	1	0.00%	50.62%
Palau	PLW	0	0.00%	18	0.00%	-
Serbia	SRB	0	0.00%	3,169	0.13%	-

DRAFT

S4. Regressions Robustness Check

This section evaluates the robustness of our main regression results along several dimensions of model specification, measurement, weighting, and outcome definition. We begin by varying the baseline return–migration models for students and workers (Section S4.1). These specifications hold fixed the core terms for gender, women's empowerment, and their interaction, while changing how we control for home-country income, time spent in the United States, and prior U.S. work or education experience. We then re-estimate the same return–migration models using country-level representativeness weights derived from DS-visa coverage (Section S4.2), allowing us to assess whether the main findings are sensitive to imbalances in the origin-country composition of the analytical sample. Together, these first checks show that the key gender–by–empowerment patterns are not driven by a particular choice of controls, functional form for GDP, or sample representativeness across origin countries.

Next, we probe the construction of the women's–empowerment variable itself. Section S4.3 replaces our preferred empowerment index with a battery of alternative indicators from UNDP and V–Dem, using otherwise identical specifications for students and workers. Section S4.4 then holds the index fixed but varies its timing, comparing models that measure empowerment at the year of first U.S. entry to models that measure it in the last observed year (the decision year). Together, these exercises test whether our conclusions depend on the particular source of the empowerment index or on when it is sampled relative to the migration spell.

We also assess the sensitivity of our results to the way gender is coded and to the choice of outcome. Section S4.5 re-estimates the return–migration models under four reasonable probability thresholds for assigning the *Female* indicator, using the gender probabilities in the resume data. Section S4.6 extends the analysis to a related educational outcome—enrollment in a U.S. PhD program—using the same structure of cohort-specific logistic models and control sets. Finally, Section S4.7 turns to scientific productivity. It documents the distributions of total and average yearly publications, reports extended OLS models for these log-transformed outcomes across migration cohorts, and visualizes the implied gender gaps as a function of women's empowerment. Across all subsections, the core message is that the qualitative patterns linking gender, women's empowerment, and migration or productivity outcomes are stable across a wide range of reasonable modeling choices.

S4.1. Return migration and women's empowerment across different models. Table S7 reports the full set of logistic regressions used to analyze return migration by cohort. Columns 1–3 present estimates for individuals who first entered the United States as students under progressively richer sets of controls, and columns 4–6 repeat the same sequence for those who first entered as workers. Across columns, the models hold fixed the core gender and women's–empowerment terms while varying the treatment of home-country income (no GDP term, residualized log GDP per capita, or raw log GDP per capita) and the inclusion of years spent in the United States and prior U.S. job or education experience. All specifications include year and origin–region fixed effects, and standard errors are clustered at the country-of-origin level.

Table S7. Rates of return migration by cohort's women's empowerment: alternative model specifications. Logistic regressions estimate the probability of returning to the origin country for migrants who first arrived as students (columns 1–3) or as workers (columns 4–6). Each specification includes indicators for gender, women's empowerment (z-score), their interaction, residualized log **GDP per capita, years spent in the United States, and prior U.S. job or education experience, with year and origin–region fixed effects. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions. Standard errors are clustered by country of origin. *Note:* Because residualized GDP is constructed by orthogonalizing log GDP with respect to women's empowerment while retaining women's empowerment in the regression, the GDP coefficient is algebraically identical across raw and residualized specifications; the difference lies in the empowerment coefficient.**

	Dependent Variable: Returned											
	Students						*Workers*					
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Female	-0.127***	-0.145***	-0.124***	-0.124***	-0.073**	-0.073**	-0.029	-0.076*	-0.044	-0.044	-0.074*	-0.074*
	(0.033)	(0.031)	(0.029)	(0.029)	(0.027)	(0.027)	(0.050)	(0.041)	(0.038)	(0.038)	(0.041)	(0.041)
Women's Empowerment (z-score)		0.193	-0.081	0.275***	-0.109	0.159*		0.269*	0.001	0.398***	-0.009	0.286**
		(0.147)	(0.098)	(0.076)	(0.101)	(0.067)		(0.118)	(0.089)	(0.102)	(0.091)	(0.098)
Female*Women's Empowerment (z-score)		0.075**	0.052*	0.052*	0.052*	0.052*		0.146**	0.136***	0.136***	0.129**	0.129**
		(0.027)	(0.023)	(0.023)	(0.026)	(0.026)		(0.049)	(0.039)	(0.039)	(0.044)	(0.044)
GDP per Capita (log)			0.437***		0.329***				0.381***		0.283***	
			(0.068)		(0.066)				(0.084)		(0.086)	
GDP per Capita (log) (Residualized)				0.437***		0.329***				0.381***		0.283***
				(0.068)		(0.066)				(0.084)		(0.086)
Years in the U.S.					-0.233***	-0.233***					-0.319***	-0.319***
					(0.012)	(0.012)					(0.016)	(0.016)
U.S. Job Experience					-0.640***	-0.640***						
					(0.055)	(0.055)						
U.S. Education Experience											-0.613***	-0.613***
											(0.074)	(0.074)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.060	0.063	0.079	0.079	0.230	0.230	0.069	0.076	0.086	0.086	0.258	0.258
Adjusted Pseudo R^2	0.060	0.063	0.078	0.078	0.229	0.229	0.069	0.076	0.086	0.086	0.258	0.258
Number of Observations	1,051,090	1,051,090	1,051,090	1,051,090	1,051,090	1,051,090	1,486,000	1,486,000	1,486,000	1,486,000	1,486,000	1,486,000
Significance levels: $^{*}p < 0.1, ^{**}p < 0.01, ^{***}p < 0.001$												

DRAFT

S4.2. Return migration with representativeness weights. Table S8 re-estimates the baseline return–migration models using the country-level representativeness weights introduced in Section S3. These weights adjust the analytical sample so that the origin-country composition more closely matches the distribution of DS-visa coverage summarized in Table S6. As in Table S7, columns 1–3 report estimates for individuals who first entered the United States as students, and columns 4–6 report the corresponding models for those who first entered as workers. Across columns, the models keep fixed the core terms for gender, women's empowerment, and their interaction, while varying the treatment of home-country income and the inclusion of years spent in the United States and prior U.S. job or education experience. All specifications include year and origin–region fixed effects, and standard errors are clustered at the country-of-origin level.

The weighted results are substantively similar to the unweighted estimates in Table S7. In the richer specifications, the coefficient on *Female* remains negative for both students and workers, indicating lower return probabilities for women conditional on the included controls. The interaction between *Female* and women's empowerment remains positive in sign but modest in magnitude and statistically imprecise, indicating that the weighted estimates do not materially alter the main conclusion that the sign of the moderation is unchanged, although the weighted interaction is estimated less precisely than in the baseline models. Likewise, the coefficients on years spent in the United States remain negative and precisely estimated, and the coefficients on prior U.S. job or education experience remain negative in the expected direction. Overall, the weighted analysis indicates that the main return–migration patterns are not driven by imbalances in the origin-country composition of the sample.

Table S8. Rates of return migration by cohort's women's empowerment: weighted alternative model specifications. Logistic regressions estimate the probability of returning to the origin country for migrants who first arrived as students (columns 1–3) or as workers (columns 4–6). Each specification includes indicators for gender, women's empowerment (z-score), their interaction, residualized log **GDP per capita, years spent in the United States, and prior U.S. job or education experience, with year and origin–region fixed effects. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions. Observations are weighted using country-specific post-stratification weights derived from the ratio of each origin country's share in DS-visa records to its share in the analytical sample, as described in Section S3 and Table S6. Standard errors are clustered by country of origin. *Note:* Because residualized GDP is constructed by orthogonalizing log GDP with respect to women's empowerment while retaining women's empowerment in the regression, the GDP coefficient is algebraically identical across raw and residualized specifications; the difference lies in the empowerment coefficient.**

	Dependent Variable: Returned											
	Students						*Workers*					
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Female	-0.143***	-0.142***	-0.128***	-0.128***	-0.099***	-0.099***	-0.132**	-0.115*	-0.106*	-0.106*	-0.141*	-0.141*
	(0.032)	(0.032)	(0.031)	(0.031)	(0.024)	(0.024)	(0.048)	(0.050)	(0.047)	(0.047)	(0.055)	(0.055)
Women's Empowerment (z-score)		0.150	-0.092	0.056	-0.101	0.001		0.032	-0.058	0.058	-0.053	0.014
		(0.113)	(0.092)	(0.080)	(0.086)	(0.068)		(0.121)	(0.107)	(0.140)	(0.098)	(0.117)
Female*Women's Empowerment (z-score)		0.023	0.009	0.009	0.017	0.017		0.060	0.061	0.061	0.057	0.057
		(0.021)	(0.019)	(0.019)	(0.022)	(0.022)		(0.043)	(0.041)	(0.041)	(0.044)	(0.044)
GDP per Capita (log)			0.336***		0.233***				0.127		0.075	
			(0.080)		(0.068)				(0.101)		(0.099)	
GDP per Capita (log) (Residualized)				0.336***		0.233***				0.127		0.075
				(0.080)		(0.068)				(0.101)		(0.099)
Years in the U.S.					-0.271***	-0.271***					-0.320***	-0.320***
					(0.017)	(0.017)					(0.024)	(0.024)
U.S. Job Experience					-0.499***	-0.499***						
					(0.083)	(0.083)						
U.S. Education Experience											-0.637***	-0.637***
											(0.104)	(0.104)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.145	0.147	0.153	0.153	0.303	0.303	0.170	0.171	0.170	0.170	0.323	0.323
Adjusted Pseudo R^2	0.145	0.147	0.153	0.153	0.303	0.303	0.170	0.171	0.170	0.170	0.323	0.323
Number of Observations	1,049,517	1,049,517	1,049,517	1,049,517	1,049,517	1,049,517	1,484,292	1,484,292	1,484,292	1,484,292	1,484,292	1,484,292
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$												

DRAFT

S4.3. Robustness to the measurement of women's empowerment. Table S9 and Table S10 examine the robustness of our return–migration models to alternative measures of women's empowerment. For students (Table S9) and workers (Table S10) separately, each column replaces the baseline women's empowerment index with a different indicator drawn from UNDP- and V-Dem–based measures (and related variants), while holding the rest of the specification fixed. Across columns, the covariate set is kept constant: all models include indicators for gender, the chosen empowerment measure, their interaction, residualized log GDP per capita, years in the United States, and U.S. job or education experience where applicable, as well as year and origin–region fixed effects with standard errors clustered by country of origin. These tables therefore allow us to assess whether the main patterns in the return–migration analysis are sensitive to the particular operationalization of women's empowerment.

Table S9. Return migration and alternative women's empowerment measures (Students). Logistic regressions estimate the probability of returning to the home country for migrants who first entered the United States as students. Each column replaces the baseline women's empowerment index (z-score) with an alternative indicator (UNDP- and V-Dem–based measures and related variants), while keeping the core specification fixed. All models include indicators for gender, the women's empowerment measure, their interaction, residualized log GDP per capita, years in the United States, and U.S. job experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Returned																				
	UNDP							*V-Dem*												
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**	**Model 13**	**Model 14**	**Model 15**	**Model 16**	**Model 17**	**Model 18**	**Model 19**	**Model 20**
Female	-0.071*	-0.072**	-0.085***	-0.082***	-0.072**	-0.084***	-0.088***	-0.072**	-0.071*	-0.069*	-0.070*	-0.074**	-0.076**	-0.071**	-0.071**	-0.074**	-0.071**	-0.073**	-0.083**	-0.073**
	(0.028)	(0.027)	(0.025)	(0.024)	(0.025)	(0.024)	(0.020)	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)	(0.026)	(0.026)	(0.027)	(0.025)	(0.027)	(0.026)	(0.026)	(0.027)
GDP per Capita (log) (Residualized)	0.261**	0.239**	0.279***	0.297***	0.325***	0.262***	0.349***	0.327***	0.325***	0.318***	0.310***	0.284***	0.274**	0.323***	0.299***	0.318***	0.288***	0.301***	0.285***	0.272***
	(0.084)	(0.088)	(0.048)	(0.048)	(0.051)	(0.056)	(0.078)	(0.066)	(0.071)	(0.064)	(0.059)	(0.068)	(0.091)	(0.052)	(0.071)	(0.051)	(0.048)	(0.052)	(0.047)	(0.062)
Years in the U.S.	-0.239***	-0.238***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.239***	-0.238***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.013)
U.S. Job Experience	-0.630***	-0.630***	-0.628***	-0.631***	-0.632***	-0.627***	-0.629***	-0.630***	-0.630***	-0.630***	-0.631***	-0.630***	-0.630***	-0.629***	-0.631***	-0.629***	-0.629***	-0.630***	-0.629***	-0.632***
	(0.057)	(0.057)	(0.056)	(0.058)	(0.057)	(0.058)	(0.057)	(0.057)	(0.056)	(0.056)	(0.057)	(0.057)	(0.057)	(0.057)	(0.057)	(0.058)	(0.057)	(0.058)	(0.056)	(0.057)
Polyarchy (z-score)	0.213***																			
	(0.052)																			
Female × Polyarchy (z-score)	0.034																			
	(0.025)																			
Liberal Democracy (z-score)		0.258***																		
		(0.043)																		
Female × Liberal Democracy (z-score)		0.043*																		
		(0.025)																		
Parliament Seat Share Gap (M-F) (z-score)			-0.254***																	
			(0.061)																	
Female × Parliament Seat Share Gap (M-F) (z-score)			-0.084***																	
			(0.020)																	
Labor Force Participation % Gap (M-F) (z-score)				-0.259**																
				(0.095)																
Female × Labor Force Participation % Gap (M-F) (z-score)				-0.067***																
				(0.017)																
Secondary Education Share Gap (M-F) (z-score)					-0.202**															
					(0.066)															
Female × Secondary Education Share Gap (M-F) (z-score)					-0.037															
					(0.033)															
World Bank Pay Score (z-score)						0.369***														
						(0.058)														
Female × World Bank Pay Score (z-score)						0.067**														
						(0.021)														
Gender Inequality Index (z-score)							-0.256***													
							(0.077)													
Female × Gender Inequality Index (z-score)							-0.103***													
							(0.020)													
Women's Empowerment (z-score)								0.155*												
								(0.068)												
Female × Women's Empowerment (z-score)								0.059*												
								(0.025)												
Women's Civil Liberties (z-score)									0.174**											
									(0.062)											
Female × Women's Civil Liberties (z-score)									0.055*											
									(0.023)											
Domestic Movement for Women (z-score)										0.128*										
										(0.067)										
Female × Domestic Movement for Women (z-score)										0.039*										
										(0.019)										
Freedom from Forced Labor for Women (z-score)											0.127*									
											(0.071)									
Female × Freedom from Forced Labor for Women (z-score)											0.060**									
											(0.023)									
Property Rights for Women (z-score)												0.274***								
												(0.060)								
Female × Property Rights for Women (z-score)												0.054**								
												(0.019)								
Access to Justice for Women (z-score)													0.268***							
													(0.051)							
Female × Access to Justice for Women (z-score)													0.078**							
													(0.024)							
Women's Civil Society Participation (z-score)														0.138*						
														(0.062)						
Female × Women's Civil Society Participation (z-score)														0.053*						
														(0.025)						
Political Discussion Freedom for Women (z-score)															0.200***					
															(0.049)					
Female × Political Discussion Freedom for Women (z-score)															0.044*					
															(0.023)					
No CSOs Participation for Women (z-score)																0.145**				
																(0.055)				
Female × No CSOs Participation for Women (z-score)																0.065**				
																(0.024)				
Female Journalists (%) (z-score)																	0.089			
																	(0.087)			
Female × Female Journalists (%) (z-score)																	0.032			
																	(0.032)			
Women's Formal Political Participation (z-score)																		0.102		
																		(0.065)		
Female × Women's Formal Political Participation (z-score)																		0.056*		
																		(0.022)		
Chamber of the Legislature (%) (z-score)																			0.236***	
																			(0.062)	
Female × Chamber of the Legislature (%) (z-score)																			0.076***	
																			(0.020)	
Political Power Equality (z-score)																				0.216**
																				(0.066)
Female × Political Power Equality (z-score)																				0.053***
																				(0.016)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231
Adjusted Pseudo R^2	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231	0.231
Number of Observations	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049	1,017,049

Significance levels: *$p < 0.1$, **$p < 0.01$, ***$p < 0.001$

DRAFT

Table S10. Return migration and alternative women's empowerment measures (Workers). Logistic regressions estimate the probability of returning to the home country for migrants who first entered the United States as workers. Each column replaces the baseline women's empowerment index (z-score) with an alternative indicator (UNDP- and V-Dem–based measures and related variants), while keeping the core specification fixed. All models include indicators for gender, the women's empowerment measure, their interaction, residualized log GDP per capita, years in the United States, and U.S. education experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

	Dependent Variable: Returned																			
	UNDP							*V-Dem*												
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**	**Model 11**	**Model 12**	**Model 13**	**Model 14**	**Model 15**	**Model 16**	**Model 17**	**Model 18**	**Model 19**	**Model 20**
Female	-0.051	-0.057	-0.076	-0.094**	-0.061	-0.090**	-0.094**	-0.071*	-0.051	-0.046	-0.047	-0.050	-0.063	-0.083*	-0.057	-0.084**	-0.054	-0.067	-0.074	-0.056
	(0.046)	(0.044)	(0.049)	(0.029)	(0.039)	(0.031)	(0.033)	(0.041)	(0.046)	(0.047)	(0.049)	(0.049)	(0.041)	(0.034)	(0.044)	(0.031)	(0.056)	(0.046)	(0.047)	(0.048)
GDP per Capita (log) (Residualized)	0.270**	0.236*	0.283**	0.255**	0.328***	0.266**	0.186	0.289**	0.359***	0.355***	0.347***	0.363***	0.280**	0.276**	0.278**	0.273**	0.310***	0.272**	0.291***	0.279**
	(0.093)	(0.101)	(0.088)	(0.080)	(0.075)	(0.093)	(0.131)	(0.091)	(0.078)	(0.077)	(0.079)	(0.075)	(0.097)	(0.091)	(0.093)	(0.088)	(0.081)	(0.091)	(0.087)	(0.092)
Years in the U.S.	-0.325***	-0.324***	-0.325***	-0.325***	-0.325***	-0.325***	-0.324***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***	-0.325***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
U.S. Education Experience	-0.615***	-0.615***	-0.612***	-0.613***	-0.620***	-0.609***	-0.618***	-0.616***	-0.619***	-0.620***	-0.614***	-0.615***	-0.617***	-0.611***	-0.616***	-0.611***	-0.616***	-0.608***	-0.613***	-0.615***
	(0.075)	(0.074)	(0.074)	(0.073)	(0.075)	(0.076)	(0.075)	(0.075)	(0.074)	(0.074)	(0.073)	(0.073)	(0.074)	(0.075)	(0.075)	(0.075)	(0.074)	(0.074)	(0.074)	(0.074)
Polyarchy (z-score)	0.230**																			
	(0.077)																			
Female × Polyarchy (z-score)	0.082*																			
	(0.036)																			
Liberal Democracy (z-score)		0.269***																		
		(0.082)																		
Female × Liberal Democracy (z-score)		0.097**																		
		(0.035)																		
Parliament Seat Share Gap (M-F) (z-score)			-0.254***																	
			(0.069)																	
Female × Parliament Seat Share Gap (M-F) (z-score)			-0.131***																	
			(0.036)																	
Labor Force Participation % Gap (M-F) (z-score)				-0.486***																
				(0.086)																
Female × Labor Force Participation % Gap (M-F) (z-score)				-0.180***																
				(0.023)																
Secondary Education Share Gap (M-F) (z-score)					-0.163															
					(0.104)															
Female × Secondary Education Share Gap (M-F) (z-score)					-0.121*															
					(0.054)															
World Bank Pay Score (z-score)						0.364**														
						(0.112)														
Female × World Bank Pay Score (z-score)						0.176***														
						(0.033)														
Gender Inequality Index (z-score)							-0.454***													
							(0.105)													
Female × Gender Inequality Index (z-score)							-0.173***													
							(0.035)													
Women's Empowerment (z-score)								0.282**												
								(0.098)												
Female × Women's Empowerment (z-score)								0.133**												
								(0.043)												
Women's Civil Liberties (z-score)									0.220**											
									(0.078)											
Female × Women's Civil Liberties (z-score)									0.102*											
									(0.041)											
Domestic Movement for Women (z-score)										0.167*										
										(0.067)										
Female × Domestic Movement for Women (z-score)										0.084*										
										(0.036)										
Freedom from Forced Labor for Women (z-score)											0.162*									
											(0.075)									
Female × Freedom from Forced Labor for Women (z-score)											0.084**									
											(0.027)									
Property Rights for Women (z-score)												0.218*								
												(0.091)								
Female × Property Rights for Women (z-score)												0.082*								
												(0.035)								
Access to Justice for Women (z-score)													0.283**							
													(0.090)							
Female × Access to Justice for Women (z-score)													0.129***							
													(0.034)							
Women's Civil Society Participation (z-score)														0.306***						
														(0.092)						
Female × Women's Civil Society Participation (z-score)														0.157***						
														(0.041)						
Political Discussion Freedom for Women (z-score)															0.250***					
															(0.074)					
Female × Political Discussion Freedom for Women (z-score)															0.100**					
															(0.037)					
No CSOs Participation for Women (z-score)																0.303***				
																(0.091)				
Female × No CSOs Participation for Women (z-score)																0.162***				
																(0.034)				
Female Journalists (%) (z-score)																	0.147*			
																	(0.079)			
Female × Female Journalists (%) (z-score)																	0.071			
																	(0.066)			
Women's Formal Political Participation (z-score)																		0.262***		
																		(0.067)		
Female × Women's Formal Political Participation (z-score)																		0.115*		
																		(0.050)		
Chamber of the Legislature (%) (z-score)																			0.238***	
																			(0.067)	
Female × Chamber of the Legislature (%) (z-score)																			0.125***	
																			(0.036)	
Political Power Equality (z-score)																				0.228**
																				(0.073)
Female × Political Power Equality (z-score)																				0.091***
																				(0.024)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.260	0.260	0.260	0.261	0.260	0.261	0.261	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260
Adjusted Pseudo R^2	0.260	0.260	0.260	0.261	0.260	0.261	0.261	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260
Number of Observations	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896	1,441,896

Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$

DRAFT

S4.4. Timing of the women's–empowerment measure and robustness. Table S11 evaluates how our return–migration results depend on the timing of the women's empowerment measure. For each cohort (students and workers), we re-estimate the baseline logistic specification twice: once using women's empowerment in the year of first U.S. entry ("First") and once using empowerment in the individual's last observed year ("Last" / decision year). Across columns, the set of covariates is held fixed—gender, the women's–empowerment index (z-score), their interaction, residualized log GDP per capita, years in the United States, and U.S. job or education experience—along with year and origin–region fixed effects and clustered standard errors at the country-of-origin level. Comparing the "First" and "Last" columns within each cohort thus isolates sensitivity to the timing convention used to operationalize women's empowerment.

Table S11. Return migration and timing of women's empowerment measurement. Logistic regressions estimate the probability of returning to the home country for migrants who first arrived as students or as workers, using women's empowerment measured either in the immigration year ("First") or in the last observed year ("Last"/ decision year). Each column reports the baseline specification with indicators for gender, women's empowerment (z-score), their interaction, residualized log GDP per capita, years in the United States, and U.S. job or education experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Returned				
	Students		*Workers*	
	Immigration Year	**Last Action Year**	**Immigration Year**	**Last Action Year**
Female	-0.073**	-0.133***	-0.074*	-0.112**
	(0.027)	(0.025)	(0.041)	(0.037)
Women's Empowerment (z-score)	0.159*	0.242***	0.286**	0.361***
	(0.067)	(0.070)	(0.098)	(0.079)
Female×Women's Empowerment (z-score)	0.052*	0.059*	0.129**	0.128**
	(0.026)	(0.023)	(0.044)	(0.040)
GDP per Capita (log) (Residualized)	0.329***	0.404***	0.283***	0.434***
	(0.066)	(0.067)	(0.086)	(0.064)
Years in the U.S.	-0.233***	-0.127***	-0.319***	-0.149***
	(0.012)	(0.007)	(0.016)	(0.010)
U.S. Job Experience	-0.640***	-0.723***		
	(0.055)	(0.058)		
U.S. Education Experience			-0.613***	-0.667***
			(0.074)	(0.087)
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.230	0.166	0.258	0.163
Adjusted Pseudo R^2	0.229	0.166	0.258	0.163
Number of Observations	1,051,090	1,049,843	1,486,000	1,481,413
Significance levels: $^*p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$				

DRAFT

S4.5. Choice of gender threshold. Table S12 examines the sensitivity of our return–migration models to the construction of the *Female* indicator. Using the gender probabilities provided in the resume data, we define *Female* under four alternative decision thresholds $T \in \{0.4, 0.5, 0.6, 0.7\}$ and, for each threshold, re-estimate the baseline logistic specification separately for students and workers. For each combination of cohort and threshold, the table reports the estimated coefficients and standard errors for *Female*, Women's Empowerment (z-score), their interaction, and residualized log GDP per capita, holding fixed the remaining controls (years in the United States, U.S. job or education experience), year and origin–region fixed effects, and clustered standard errors at the country-of-origin level. This setup allows us to assess how the key gender and empowerment terms respond to reasonable changes in the classification rule used to assign gender.

Table S12. Sensitivity to the gender-classification threshold. Logistic regressions estimate the probability of returning to the home country for migrants who first arrived as students or as workers. Columns vary the threshold T used to classify an observation as female, with $T \in \{0.4, 0.5, 0.6, 0.7\}$. Each column reports the baseline specification with indicators for gender, women's empowerment (z-score), their interaction, residualized log GDP per capita, years in the United States, and U.S. job or education experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Returned								
	Students				*Workers*			
	T=0.4	**T=0.5**	**T=0.6**	**T=0.7**	**T=0.4**	**T=0.5**	**T=0.6**	**T=0.7**
Female	-0.080**	-0.073**	-0.074**	-0.070*	-0.075*	-0.074*	-0.075*	-0.072*
	(0.026)	(0.027)	(0.027)	(0.028)	(0.040)	(0.041)	(0.042)	(0.042)
Women's Empowerment (z-score)	0.157*	0.159*	0.159*	0.159*	0.285**	0.286**	0.286**	0.287**
	(0.067)	(0.067)	(0.068)	(0.068)	(0.099)	(0.098)	(0.097)	(0.097)
Female×Women's Empowerment (z-score)	0.054*	0.052*	0.055*	0.055*	0.127**	0.129**	0.130**	0.131**
	(0.026)	(0.026)	(0.026)	(0.027)	(0.042)	(0.044)	(0.044)	(0.045)
GDP per Capita (log) (Residualized)	0.329***	0.329***	0.329***	0.330***	0.283***	0.283***	0.283***	0.283***
	(0.065)	(0.066)	(0.065)	(0.065)	(0.086)	(0.086)	(0.086)	(0.085)
Years in the U.S.	-0.233***	-0.233***	-0.233***	-0.233***	-0.319***	-0.319***	-0.319***	-0.319***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)	(0.016)	(0.016)	(0.016)
U.S. Job Experience	-0.640***	-0.640***	-0.640***	-0.640***				
	(0.055)	(0.055)	(0.055)	(0.055)				
U.S. Education Experience					-0.612***	-0.613***	-0.613***	-0.613***
					(0.074)	(0.074)	(0.074)	(0.074)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.230	0.230	0.230	0.230	0.258	0.258	0.258	0.258
Adjusted Pseudo R^2	0.230	0.229	0.229	0.229	0.258	0.258	0.258	0.258
Number of Observations	1,051,090	1,051,090	1,051,090	1,051,090	1,486,000	1,486,000	1,486,000	1,486,000
Significance levels: $^*p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$								

DRAFT

S4.6. PhD Enrollment. Table S13 presents the full set of logistic regressions used to analyze enrollment in a U.S. PhD program. Columns 1–3 report estimates for individuals who first entered the United States as students, and columns 4–6 report the corresponding models for those who first entered as workers. Across columns, the models hold fixed the core terms for gender, women's empowerment (z-score), and their interaction, while varying the set of controls, including home-country income, time spent in the United States, and prior U.S. work or education experience. All specifications include year and origin–region fixed effects, with standard errors clustered at the country-of-origin level, providing a consistent framework for comparing PhD enrollment patterns across entry cohorts and control sets.

Table S14 adds an additional robustness check focused on the student cohort. Here we re-estimate the student models after excluding individuals whose first U.S. education record is already a PhD program, that is, those who migrated directly for doctoral study. This restriction isolates the population of students who face the decision to start a PhD after arrival in the United States (for example, after a bachelor's or master's degree), and therefore removes any potential bias arising from selection into direct PhD migration. The specification mirrors the sequence in Table S13: we progressively add controls for home-country income, duration in the United States, and prior U.S. job experience while keeping the gender and women's–empowerment terms fixed. The resulting coefficients are similar in sign and magnitude to those in the full student sample, indicating that our main conclusions about gender, women's empowerment, and PhD enrollment are not driven by individuals who arrive in the United States specifically for a PhD program.

Finally, Table S15 replaces enrollment in a U.S. PhD program with earning a U.S. PhD degree as the outcome. The results remain similar, indicating that our conclusions are robust to whether we measure doctoral participation through enrollment or completed attainment.

Table S13. U.S. PhD enrollment and women's empowerment: alternative model specifications. Logistic regressions estimate the probability of enrolling in a U.S. PhD program for migrants who first arrived as students (columns 1–4) and as workers (columns 5–8). Across columns, we hold fixed the core terms for gender, women's empowerment (z-score), and their interaction, while varying the set of controls, including residualized log GDP per capita, years spent in the United States, and prior U.S. job or education experience. All models include year and origin–region fixed effects, with standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Enrolling in a U.S. Ph.D. Program								
	Students				*Workers*			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Female	-0.193**	-0.181**	-0.200***	-0.237***	-0.051	-0.002	-0.015	-0.039
	(0.063)	(0.058)	(0.057)	(0.058)	(0.058)	(0.045)	(0.045)	(0.045)
Women's Empowerment (z-score)		-0.338*	-0.418***	-0.360***		-0.376***	-0.466***	-0.392***
		(0.143)	(0.099)	(0.093)		(0.098)	(0.106)	(0.091)
Female×Women's Empowerment (z-score)		-0.007	0.010	0.012		0.069*	0.068*	0.079*
		(0.039)	(0.036)	(0.037)		(0.039)	(0.039)	(0.036)
GDP per Capita (log) (Residualized)			-0.339***	-0.250***			-0.171*	-0.100
			(0.067)	(0.066)			(0.082)	(0.068)
Years in the U.S.				0.090***				0.255***
				(0.007)				(0.015)
U.S. Job Experience				0.304***				
				(0.065)				
U.S. Education Experience								0.937***
								(0.072)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.063	0.070	0.080	0.112	0.060	0.067	0.068	0.121
Adjusted Pseudo R^2	0.063	0.070	0.080	0.112	0.060	0.067	0.068	0.120
Number of Observations	1,051,090	1,051,090	1,051,090	1,051,090	1,486,000	1,486,000	1,486,000	1,486,000
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$								

DRAFT

Table S14. U.S. PhD enrollment among students excluding migrants whose first U.S. education record is a PhD. Logistic regressions estimate the probability of enrolling in a U.S. PhD program for migrants who first arrived as students, excluding those whose first observed U.S. degree is doctoral. Across columns, we hold fixed the core terms for gender, women's empowerment (z-score), and their interaction, while varying the set of controls, including GDP per capita (log or residualized log), years spent in the United States, and U.S. job experience. All models include year and origin–region fixed effects, with standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Enrolling in a U.S. Ph.D. Program				
	Model 1	**Model 2**	**Model 3**	**Model 4**
Female	-0.148*	-0.133*	-0.144**	-0.201***
	(0.061)	(0.057)	(0.055)	(0.050)
Women's Empowerment (z-score)		-0.356**	-0.413***	-0.344***
		(0.112)	(0.087)	(0.079)
Female×Women's Empowerment (z-score)		-0.001	0.014	0.025
		(0.033)	(0.032)	(0.033)
GDP per Capita (log) (Residualized)			-0.205***	-0.079
			(0.060)	(0.057)
Years in the U.S.				0.140***
				(0.013)
U.S. Job Experience				0.631***
				(0.070)
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.068	0.074	0.077	0.136
Adjusted Pseudo R^2	0.068	0.074	0.077	0.136
Number of Observations	988,791	988,791	988,791	988,791
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$				

Table S15. U.S. PhD attainment and women's empowerment by cohort. Logistic regressions estimate the probability of attaining a U.S. PhD degree for migrants who first arrived as students or as workers. Each specification includes indicators for gender, women's empowerment (z-score), their interaction, residualized log GDP per capita, years in the United States, and prior U.S. job or education experience, with year and origin–region fixed effects and standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: Attaining a U.S. Ph.D. Degree		
	Students	**Workers**
Female	-0.261***	-0.040
	(0.055)	(0.038)
Women's Empowerment (z-score)	-0.339***	-0.423***
	(0.085)	(0.085)
Female×Women's Empowerment (z-score)	0.025	0.116***
	(0.040)	(0.035)
GDP per Capita (log) (Residualized)	-0.262***	-0.126*
	(0.058)	(0.067)
Years in the U.S.	0.076***	0.183***
	(0.005)	(0.014)
U.S. Job Experience	0.291***	
	(0.057)	
U.S. Education Experience		0.756***
		(0.076)
Year FE	Yes	Yes
Region FE	Yes	Yes
Pseudo R^2	0.187	0.154
Adjusted Pseudo R^2	0.187	0.154
Number of Observations	1,051,090	1,396,907
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$		

DRAFT

S4.7. Scientific productivity: distributions and extended models. Figures S11–S14 and Tables S16–S17 provide additional detail on the construction and modeling of the publication outcomes used in the main text.

Figure S11 displays the empirical distribution of log total publications per individual, summarizing the skewness of raw publication counts and motivating the log transformation applied in the analysis, while Figure S12 shows the analogous distribution for our second outcome—log average yearly publications per individual (endpoint 2022)—using the same transformation on publication counts normalized by career length, defined as 2022 minus the year of first recorded publication.

Table S16 reports the full set of OLS regressions where the dependent variable is log total publications. Columns distinguish Immigrants, Returnees, Stayers, and Movers relative to Non-Immigrants, as well as Stayers relative to Returnees. Each specification includes gender, women's empowerment (z-score), their interactions with migration cohort, residualized log GDP per capita, and fixed effects for region and field.

Figure S13 visualizes the model-implied patterns for total publications. The top row plots marginal predicted means of log total publications for men and women across the women's empowerment z-score, and the bottom row presents the corresponding percent change in publications for women relative to men, with 95% confidence intervals, for each migration-cohort comparison reported in Table S16.

Table S17 presents the corresponding regressions for log average yearly publications. The covariate set, interaction terms, and fixed effects mirror the total-publications specifications, again distinguishing Immigrants, Returnees, Stayers, Movers, and their comparisons to Non-Immigrants and to each other.

Finally, Figure S14 provides the model-based predictions for average yearly publications. As with total publications, the top row plots marginal predicted means of log average yearly publications for men and women across the women's empowerment z-score, and the bottom row reports the percent change in women's publications relative to men, with 95% confidence intervals, for each migration-cohort comparison in Table S17.

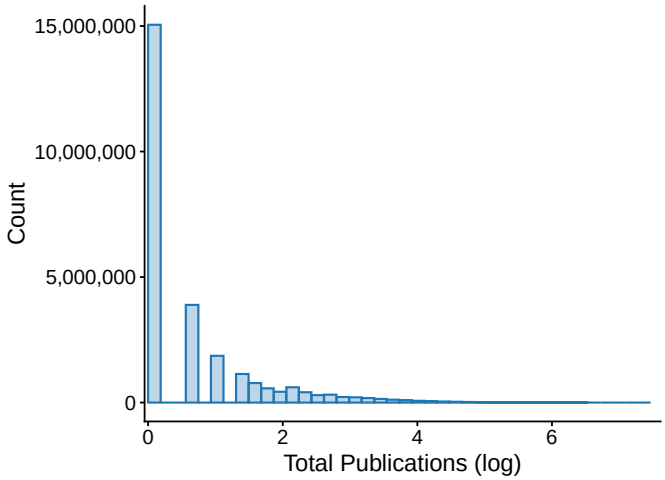


Fig. S11. Distribution of total publications. Histogram of log total publications per individual. The x-axis reports log(total publications) and the y-axis shows the number of individuals.

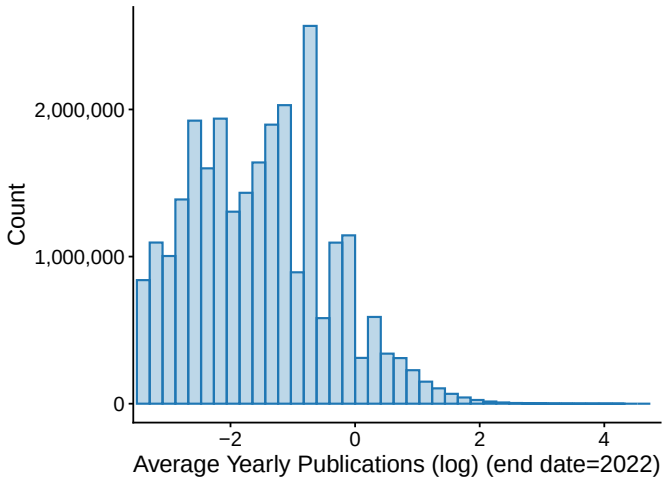


Fig. S12. Distribution of average yearly publications. Histogram of log average yearly publications per individual (endpoint 2022). The x-axis reports log(average yearly publications), where average yearly publications equal total publications divided by career length (2022 minus year of first publication); the y-axis shows the number of individuals.

Table S16. Total publications and women's empowerment by migration cohort (extended models). OLS regressions estimate log total publications as a function of gender, women's empowerment (z-score), their interaction, cohort indicators, and residualized $\log$ GDP per capita. Columns compare Immigrants, Returnees, Stayers, and Movers to Non-Immigrants, as well as Stayers to Returnees. All models include fixed effects for region and field, with standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

	Dependent Variable: log total publications				
	Immigrants vs. Non-Immigrants	Returnees vs. Non-Immigrants	Stayers vs. Non-Immigrants	Movers vs. Non-Immigrants	Stayers vs. Returnees
Female	0.085***	0.085***	0.085***	0.085***	-0.231***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.033)
Women's Empowerment (z-score)	0.053	0.053	0.053	0.053	0.104
	(0.037)	(0.037)	(0.037)	(0.037)	(0.088)
Female×Women's Empowerment (z-score)	0.064***	0.064***	0.064***	0.064***	0.053*
	(0.013)	(0.013)	(0.013)	(0.013)	(0.030)
Immigrant	0.681***				
	(0.077)				
Female×Immigrant	-0.349***				
	(0.030)				
Female×Women's Empowerment (z-score)×Immigrant	-0.000				
	(0.020)				
Returnee		0.569***			
		(0.076)			
Female×Returnee		-0.346***			
		(0.032)			
Female×Women's Empowerment (z-score)×Returnee		0.013			
		(0.022)			
Stayer			0.697***		0.140***
			(0.080)		(0.032)
Female×Stayer			-0.340***		-0.006
			(0.033)		(0.021)
Female×Women's Empowerment (z-score)×Stayer			-0.007		-0.022
			(0.023)		(0.032)
Mover				0.803***	
				(0.076)	
Female×Mover				-0.416***	
				(0.026)	
Female×Women's Empowerment (z-score)×Mover				0.030	
				(0.020)	
GDP per Capita (log) (Residualized)	-0.089*	-0.089*	-0.089*	-0.089*	-0.127**
	(0.040)	(0.040)	(0.040)	(0.040)	(0.041)
Region FE	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes
R^2	0.037	0.035	0.036	0.035	0.082
Adjusted R^2	0.037	0.035	0.036	0.035	0.081
Number of Observations	26,577,392	26,441,039	26,502,709	26,417,446	159,946
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$					

DRAFT

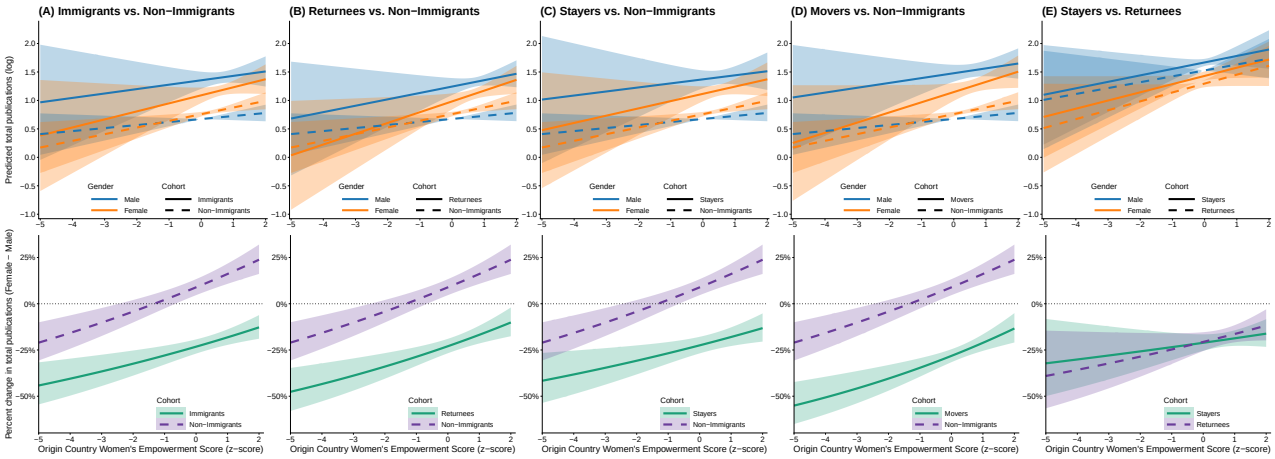


Fig. S13. Predicted total publications and gender gap by women's empowerment. The top row plots marginal predicted means of $\log$ total publications for men and women across the origin country women's empowerment (z-score); the bottom row shows the percent change in publications for women relative to men, with 95% confidence intervals. Panels correspond to cohort comparisons: **(A)** Immigrant vs. Non-Immigrant, **(B)** Returnee vs. Non-Immigrant, **(C)** Stayer vs. Non-Immigrant, **(D)** Mover vs. Non-Immigrant, and **(E)** Stayer vs. Returnee. Estimates are based on the models in Table S16, which control for region, field, and residualized $\log$ GDP per capita.

Table S17. Average yearly publications and women's empowerment by migration cohort. OLS regressions estimate log average yearly publications (endpoint 2022) as a function of gender, women's empowerment (z-score), their interaction, cohort indicators, and residualized log GDP per capita. Columns compare Immigrants, Returnees, Stayers, and Movers to Non-Immigrants, as well as Stayers to Returnees. All models include fixed effects for region and field, with standard errors clustered by country of origin. Region fixed effects account for broad geographic differences across migrants' origin countries by grouping countries into major world regions.

Dependent Variable: log average yearly publications (end date=2022)					
	Immigrants vs. Non-Immigrants	Returnees vs. Non-Immigrants	Stayers vs. Non-Immigrants	Movers vs. Non-Immigrants	Stayers vs. Returnees
Female	0.277***	0.277***	0.277***	0.277***	-0.151***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.029)
Women's Empowerment (z-score)	-0.073	-0.073	-0.073	-0.073	0.202**
	(0.062)	(0.062)	(0.062)	(0.062)	(0.064)
Female×Women's Empowerment (z-score)	0.103*	0.103*	0.103*	0.103*	0.045
	(0.040)	(0.040)	(0.040)	(0.040)	(0.031)
Immigrant	0.601***				
	(0.066)				
Female×Immigrant	-0.436***				
	(0.046)				
Female×Women's Empowerment (z-score)×Immigrant	-0.036				
	(0.040)				
Returnee		0.412***			
		(0.072)			
Female×Returnee		-0.445***			
		(0.056)			
Female×Women's Empowerment (z-score)×Returnee		-0.032			
		(0.054)			
Stayer			0.645***		0.251***
			(0.067)		(0.027)
Female×Stayer			-0.420***		0.011
			(0.048)		(0.025)
Female×Women's Empowerment (z-score)×Stayer			-0.037		-0.006
			(0.038)		(0.025)
Mover				0.722***	
				(0.067)	
Female×Mover				-0.500***	
				(0.037)	
Female×Women's Empowerment (z-score)×Mover				0.007	
				(0.032)	
GDP per Capita (log) (Residualized)	0.292*	0.292*	0.292*	0.292*	0.249**
	(0.154)	(0.155)	(0.154)	(0.155)	(0.088)
Region FE	Yes	Yes	Yes	Yes	Yes
Field FE	Yes	Yes	Yes	Yes	Yes
R^2	0.078	0.077	0.077	0.077	0.097
Adjusted R^2	0.078	0.077	0.077	0.077	0.097
Number of Observations	26,577,392	26,441,039	26,502,709	26,417,446	159,946
Significance levels: $^{*}p < 0.1$, $^{**}p < 0.01$, $^{***}p < 0.001$					

DRAFT

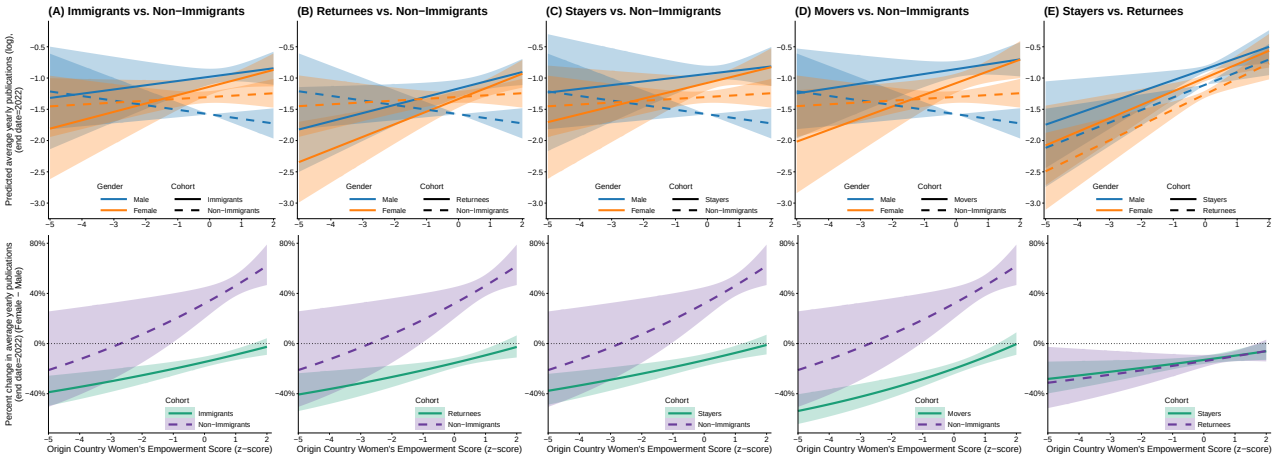


Fig. S14. Predicted average yearly publications and gender gap by women's empowerment. The top row plots marginal predicted means of log average yearly publications for men and women across the origin country women's empowerment (z-score); the bottom row shows the percent change in publications for women relative to men, with 95% confidence intervals. Panels correspond to cohort comparisons: **(A)** Immigrant vs. Non-Immigrant, **(B)** Returnee vs. Non-Immigrant, **(C)** Stayer vs. Non-Immigrant, **(D)** Mover vs. Non-Immigrant, and **(E)** Stayer vs. Returnee. Estimates are based on the models in Table S17, which control for region, field, and residualized log GDP per capita.

S5. Derivation of reported percentage effects

This section documents how the main percentage effects reported in the Introduction and Results sections are obtained from the regression models. We focus on three sets of models: (i) return migration (Table 1), (ii) U.S. PhD enrollment (Table 2), and (iii) scientific productivity (Table 3).

Throughout, we interpret logit coefficients in terms of odds ratios and log–linear coefficients in terms of multiplicative changes in publication counts. When X is a coefficient on an indicator, we report $(\exp(X) - 1) \times 100\%$ as the implied percentage difference in odds (for logistic models) or in expected publication counts (for the log–linear model).

We denote by WE the Women's Empowerment index in standardized (z–score) units, so that a one–unit change in WE corresponds to a one–standard–deviation increase.

S5.1. Return migration. The return models in Table 1 are estimated separately for migrants who first arrived as students and as workers. In each case, we estimate a logistic regression of the form

$$\log \frac{p}{1-p} = \alpha * \beta_F F * \beta_{\mathrm{WE}} \mathrm{WE} * \beta_{F\times\mathrm{WE}} F \times \mathrm{WE} * \dots$$

where F is an indicator for *Female*. We focus here on the coefficients β_F, β_{WE}, and $\beta_{F\times\mathrm{WE}}$:

- **Students:** $\beta_F = -0.073$, $\beta_{\mathrm{WE}} = 0.159$, $\beta_{F\times\mathrm{WE}} = 0.052$.
- **Workers:** $\beta_F = -0.074$, $\beta_{\mathrm{WE}} = 0.286$, $\beta_{F\times\mathrm{WE}} = 0.129$.

Female vs. male odds at mean empowerment. At $\mathrm{WE} = 0$, the female–male log–odds gap is just β_F:

$$\Delta_{\text{female vs. male}}(\mathrm{WE} = 0) = \beta_F.$$

The corresponding odds ratio is $\exp(\beta_F)$.

- Students: $\exp(-0.073) \approx 0.930$, i.e., women's odds of return are about 7.0% lower than men's.
- Workers: $\exp(-0.074) \approx 0.929$, i.e., women's odds of return are about 7.1% lower than men's.

This yields the statement that, at the average level of empowerment, women's odds of return are about 7% lower than men's in both cohorts.

Effect of a one–SD increase in empowerment. For men, a one–unit increase in WE changes the log–odds by β_{WE}, so the odds are multiplied by $\exp(\beta_{\mathrm{WE}})$. For women, the effect is $\beta_{\mathrm{WE}} + \beta_{F\times\mathrm{WE}}$.

- Students:

$$\exp(0.159) \approx 1.172 \quad (\text{men},\ +17.2\%\ \text{odds})$$
$$\exp(0.159 + 0.052) = \exp(0.211) \approx 1.235 \quad (\text{women},\ +23.5\%\ \text{odds}).$$

- Workers:

$$\exp(0.286) \approx 1.331 \quad (\text{men},\ +33.1\%\ \text{odds})$$
$$\exp(0.286 + 0.129) = \exp(0.415) \approx 1.514 \quad (\text{women},\ +51.4\%\ \text{odds}).$$

These calculations underlie the text statement that a one–standard–deviation increase in empowerment corresponds to about a 17% increase in men's odds of return and a 23% increase for women among those who first arrived as students, and to about 33% and 51% higher odds of return for men and women, respectively, among those who first arrived as workers.

Female vs. male odds across the empowerment range. At a given empowerment level $\mathrm{WE} = w$, the female–male log–odds difference is

$$\Delta(w) = \beta_F + \beta_{F\times\mathrm{WE}}, w, \quad \mathrm{OR}(w) = \exp(\Delta(w)).$$

For students ($\beta_F = -0.073$, $\beta_{F\times\mathrm{WE}} = 0.052$), this yields:

$w = -1$: $\Delta = -0.125$, $\exp(\Delta) \approx 0.882$ (–11.8%) $w = -2$: $\Delta = -0.177$, $\exp(\Delta) \approx 0.838$ (–16.2%) $w = +2$: $\Delta = 0.031$, $\exp(\Delta) \approx 1.031$ (+3.1%).

For workers ($\beta_F = -0.074$, $\beta_{F\times\mathrm{WE}} = 0.129$), this yields:

$w = -1$: $\Delta = -0.203$, $\exp(\Delta) \approx 0.816$ (–18.4%) $w = -2$: $\Delta = -0.332$, $\exp(\Delta) \approx 0.717$ (–28.3%) $w = +2$: $\Delta = 0.184$, $\exp(\Delta) \approx 1.202$ (+20.2%).

The female–male return gap therefore becomes positive at

$$w = -\frac{\beta_F}{\beta_{F\times\mathrm{WE}}},$$

which is approximately 1.40 for students and 0.57 for workers. This supports the text statement that the gap closes around the U.S. level of empowerment and becomes positive at the top of the distribution, especially among workers.

S5.2. U.S. PhD enrollment. The PhD enrollment models in Table 2 have the same interaction structure:

$$\log \frac{p}{1-p} = \alpha * \beta_F F * \beta_{\mathrm{WE}} \mathrm{WE} * \beta_{F\times\mathrm{WE}} F \times \mathrm{WE} * \dots$$

Key coefficients:

- **Students:** $\beta_F = -0.237$, $\beta_{\mathrm{WE}} = -0.360$, $\beta_{F\times\mathrm{WE}} = 0.012$.
- **Workers:** $\beta_F = -0.039$, $\beta_{\mathrm{WE}} = -0.392$, $\beta_{F\times\mathrm{WE}} = 0.079$.

DRAFT

Female vs. male odds at mean empowerment. At WE = 0:

$$\Delta(\text{WE} = 0) = \beta_F, \quad \text{OR} = \exp(\beta_F).$$

- Students: $\exp(-0.237) \approx 0.789$ (women's odds are about 21.1% lower than men's).
- Workers: $\exp(-0.039) \approx 0.962$ (women's odds are about 3.8% lower than men's).

These values underpin the text statement that women are less likely than men to enroll in a U.S. PhD among those who first arrived as students, whereas among those who first arrived as workers the average gender difference is small and not statistically distinguishable from zero.

Effect of a one–SD increase in empowerment. For men, the effect is β_{WE}; for women, $\beta_{\text{WE}} + \beta_{F\times\text{WE}}$.

- Students:

$$\exp(-0.360) \approx 0.698 \quad \text{(men, 30.2\% decline)}$$
$$\exp(-0.360 + 0.012) = \exp(-0.348) \approx 0.706 \quad \text{(women, 29.4\% decline)}.$$

- Workers:

$$\exp(-0.392) \approx 0.676 \quad \text{(men, 32.4\% decline)}$$
$$\exp(-0.392 + 0.079) = \exp(-0.313) \approx 0.731 \quad \text{(women, 26.9\% decline)}.$$

These calculations support the text statement that higher women's empowerment is associated with lower U.S. PhD enrollment overall, but that the decline is less steep for women than for men among workers.

Female vs. male odds across the empowerment range. At empowerment level WE = w, the female–male log–odds difference is

$$\Delta(w) = \beta_F + \beta_{F\times\text{WE}}, w.$$

For students:

$$\Delta_{\text{students}}(w) = -0.237 + 0.012w.$$

For workers:

$$\Delta_{\text{workers}}(w) = -0.039 + 0.079w.$$

Evaluating at representative values:

- Students:

$$\Delta(-5) = -0.297, \quad \exp(-0.297) \approx 0.743 \Rightarrow \text{women's odds are 25.7\% lower;}$$
$$\Delta(2) = -0.213, \quad \exp(-0.213) \approx 0.808 \Rightarrow \text{women's odds are 19.2\% lower.}$$

- Workers:

$$\Delta(-5) = -0.434, \quad \exp(-0.434) \approx 0.648 \Rightarrow \text{women's odds are 35.2\% lower;}$$
$$\Delta(2) = 0.119, \quad \exp(0.119) \approx 1.126 \Rightarrow \text{women's odds are 12.6\% higher.}$$

For workers, the female–male PhD enrollment gap crosses zero at

$$w = -\frac{\beta_F}{\beta_{F\times\text{WE}}} = \frac{0.039}{0.079} \approx 0.49,$$

which supports the text statement that the worker gender gap narrows sharply with empowerment and slightly reverses at the upper end of the observed range.

S5.3. Scientific productivity. The publication model in Table 3 is log–linear:

$$\log Y = \alpha * \beta_{\text{Female}} F * \beta_{\text{WE}} \text{WE} * \beta_{F\times\text{WE}} F \times \text{WE} * \beta_{\text{Cohort}} C * \beta_{F\times C} F \times C * \beta_{F\times\text{WE}\times C} F \times \text{WE} \times C * \ldots$$

where Y is total publication count, C indicates cohort, and the omitted comparison group depends on the model.

Migration effects at mean empowerment. From the main model comparisons:

- Immigrants vs. Non–Immigrants: $\beta_{\text{Immigrant}} = 0.681$.

$$\exp(0.681) \approx 1.976 \Rightarrow \text{about 97.6\% more publications.}$$

- Returnees vs. Non–Immigrants: $\beta_{\text{Returnee}} = 0.569$.

$$\exp(0.569) \approx 1.766 \Rightarrow \text{about 76.6\% more publications.}$$

- Stayers vs. Returnees: $\beta_{\text{Stayer}} = 0.140$.

$$\exp(0.140) \approx 1.150 \Rightarrow \text{about 15.0\% more publications.}$$

These calculations underlie the text statement that immigrants publish about 98% more papers than comparable non–immigrants, returnees publish about 77% more, and male stayers publish about 15% more than male returnees.

Female vs. male productivity among non–immigrants. For non–immigrants, the female–male log difference is

$$\Delta_{\text{nonimm}}(w) = \beta_{\text{Female}} + \beta_{F\times\text{WE}}, w,$$

with

$$\beta_{\text{Female}} = 0.085, \qquad \beta_{F\times\text{WE}} = 0.064.$$

- At WE = −5:

$$\Delta_{\text{nonimm}}(-5) = 0.085 + 0.064 \times (-5) = -0.235,$$
$$\exp(-0.235) \approx 0.791 \Rightarrow \text{women publish 20.9\% fewer papers than men.}$$

- At WE = 2:

$$\Delta_{\text{nonimm}}(2) = 0.085 + 0.064 \times 2 = 0.213,$$
$$\exp(0.213) \approx 1.237 \Rightarrow \text{women publish 23.7\% more papers than men.}$$

These values correspond to the text statement that non–immigrant women publish about 21% fewer papers than non–immigrant men in the least empowering settings and about 24% more in highly empowering contexts.

DRAFT

Female vs. male productivity among returnees. From the "Stayers vs. Returnees" main model, the female–male log gap among returnees is the baseline female effect:

$$\Delta_{\text{returnee}}(w) = \beta_{\text{Female}} + \beta_{F\times\text{WE}}, w,$$

with

$$\beta_{\text{Female}} = -0.231, \qquad \beta_{F\times\text{WE}} = 0.053.$$

Thus

$$\Delta_{\text{returnee}}(w) = -0.231 + 0.053w.$$

Evaluating at representative values:

- At WE $= -2$:

$$\Delta_{\text{returnee}}(-2) = -0.231 - 0.106 = -0.337,$$
$$\exp(-0.337) \approx 0.714 \Rightarrow \text{women publish 28.6\% fewer papers than men.}$$

- At WE $= 2$:

$$\Delta_{\text{returnee}}(2) = -0.231 + 0.106 = -0.125,$$
$$\exp(-0.125) \approx 0.883 \Rightarrow \text{women publish 11.7\% fewer papers than men.}$$

These values motivate the statement that, among returnees, women publish about 29% fewer papers than men in moderately low-empowerment contexts and about 12% fewer in high-empowerment contexts.

Female vs. male productivity among stayers. From the "Stayers vs. Returnees" main model, the female–male log gap among stayers is

$$\Delta_{\text{stayer}}(w) = \beta_{\text{Female}} + \beta_{F\times\text{Stayer}} * (\beta_{F\times\text{WE}} + \beta_{F\times\text{WE}\times\text{Stayer}}), w,$$

with

$$\beta_{\text{Female}} = -0.231, \quad \beta_{F\times\text{Stayer}} = -0.006, \quad \beta_{F\times\text{WE}} = 0.053, \quad \beta_{F\times\text{WE}\times\text{Stayer}} = -0.022,$$

so

$$\Delta_{\text{stayer}}(w) = -0.237 + 0.031w.$$

- At WE $= -2$:

$$\Delta_{\text{stayer}}(-2) = -0.237 - 0.062 = -0.299,$$
$$\exp(-0.299) \approx 0.742 \Rightarrow \text{women publish 25.8\% fewer papers than men.}$$

- At WE $= 2$:

$$\Delta_{\text{stayer}}(2) = -0.237 + 0.062 = -0.175,$$
$$\exp(-0.175) \approx 0.840 \Rightarrow \text{women publish 16.0\% fewer papers than men.}$$

These values support the statement that, among stayers, the female shortfall declines from about 26% in moderately low-empowerment settings to about 16% in high-empowerment settings.

Female stayers vs. female returnees. The log difference between female stayers and female returnees comes from the "Stayers vs. Returnees" model:

$$\Delta_{\text{female, stay vs. return}}(w) = \beta_{\text{Stayer}} + \beta_{F\times\text{Stayer}} + \beta_{F\times\text{WE}\times\text{Stayer}}, w,$$

with

$$\beta_{\text{Stayer}} = 0.140, \quad \beta_{F\times\text{Stayer}} = -0.006, \quad \beta_{F\times\text{WE}\times\text{Stayer}} = -0.022.$$

Thus

$$\Delta_{\text{female, stay vs. return}}(w) = 0.134 - 0.022w.$$

- At WE $= -5$:

$$\Delta(-5) = 0.134 + 0.110 = 0.244,$$
$$\exp(0.244) \approx 1.276 \Rightarrow \text{female stayers publish 27.6\% more than female returnees.}$$

- At WE $= 0$:

$$\Delta(0) = 0.134,$$
$$\exp(0.134) \approx 1.143 \Rightarrow \text{female stayers publish 14.3\% more than female returnees.}$$

- At WE $= 2$:

$$\Delta(2) = 0.134 - 0.044 = 0.090,$$
$$\exp(0.090) \approx 1.094 \Rightarrow \text{female stayers publish 9.4\% more than female returnees.}$$

These calculations underlie the statement that female stayers publish about 14% more than female returnees at the mean level of women's empowerment, and that this premium varies from about 28% in the lowest-empowerment contexts to about 9% in the highest.

S5.4. Summary of reported percentage effects. Table S18 summarizes the main percentage effects reported in the text, the approximate values used in the narrative, and the underlying calculations from the regression models.

DRAFT

Table S18. Summary of reported percentage effects and their derivation from regression coefficients.

Location	Quantity described in text	Reported value	Exact calculation
Introduction / Return migration	Women's odds of return vs. men at mean empowerment	$\approx 7\%$ lower	Students: $\exp(-0.073) - 1 = -0.070$; Workers: $\exp(-0.074) - 1 = -0.071$ (Table 1, β_F).
Introduction / Return migration	Female odds of return vs. men in low-empowerment settings (1–2 SD below mean)	12–28% lower	Students: $\exp(-0.073 + 0.052 \times (-1)) - 1 = -0.118$; Workers: $\exp(-0.074 + 0.129 \times (-2)) - 1 = -0.283$.
Introduction / Return migration	Female odds of return vs. men in high-empowerment settings (top of distribution)	3–20% higher	Students: $\exp(-0.073+0.052\times 2)-1 = 0.031$; Workers: $\exp(-0.074+0.129 \times 2) - 1 = 0.202$.
Return migration	Effect of one–SD increase in empowerment on odds of return	Men: 17%, 33% higher; Women: 24%, 51% higher	Students: men $\exp(0.159)-1 = 0.172$, women $\exp(0.159+0.052)-1 = 0.235$; Workers: men $\exp(0.286) - 1 = 0.331$, women $\exp(0.286 + 0.129) - 1 = 0.514$.
Introduction / PhD enrollment	Women's odds of U.S. PhD enrollment vs. men at mean empowerment	21% lower for students; $\approx 4\%$ lower for workers	Students: $\exp(-0.237) - 1 = -0.211$; Workers: $\exp(-0.039) - 1 = -0.038$ (Table 2, β_F).
Introduction / PhD enrollment	Effect of one–SD increase in empowerment on PhD enrollment odds	Students: 30% and 29% decline; Workers: 32% and 27% decline	Students: men $\exp(-0.360) - 1 = -0.302$, women $\exp(-0.360 + 0.012) - 1 = -0.294$; Workers: men $\exp(-0.392) - 1 = -0.324$, women $\exp(-0.392 + 0.079) - 1 = -0.269$.
Introduction / Scientific productivity	Immigrants vs. non–immigrants: total publications	$\approx 98\%$ more	$\exp(0.681) - 1 = 0.976$ (Table 3, $\beta_{\text{Immigrant}}$).
Introduction / Scientific productivity	Returnees vs. non–immigrants: total publications	$\approx 77\%$ more	$\exp(0.569) - 1 = 0.766$ (Table 3, β_{Returnee}).
Scientific productivity	Male stayers vs. male returnees: total publications	$\approx 15\%$ more	$\exp(0.140) - 1 = 0.150$ (Table 3, β_{Stayer} in the Stayers vs. Returnees model).
Scientific productivity	Non–immigrant women vs. men: low vs. high empowerment	$\approx$ 21% fewer to 24% more	Female–male gap: $\Delta_{\text{nonimm}}(w) = 0.085 + 0.064w$. At WE $= -5$: $\exp(-0.235) - 1 = -0.209$; at WE $= 2$: $\exp(0.213) - 1 = 0.237$.
Scientific productivity	Returnee women vs. men: moderately low vs. high empowerment	$\approx 29\%$ fewer to $\approx 12\%$ fewer	$\Delta_{\text{returnee}}(w) = -0.231 + 0.053w$. At WE $= -2$: $\exp(-0.337) - 1 = -0.286$; at WE $= 2$: $\exp(-0.125) - 1 = -0.117$.
Scientific productivity	Stayer women vs. men: moderately low vs. high empowerment	$\approx 26\%$ to 16% fewer	$\Delta_{\text{stayer}}(w) = -0.237 + 0.031w$. At WE $= -2$: $\exp(-0.299) - 1 = -0.258$; at WE $= 2$: $\exp(-0.175) - 1 = -0.160$.
Scientific productivity	Female stayers vs. female returnees: low, mean, and high empowerment	$\approx$ 28%, 14%, and 9% more	$\Delta_{\text{female, stay vs. return}}(w) = 0.134 - 0.022w$. At WE $= -5$: $\exp(0.244) - 1 = 0.276$; at WE $= 0$: $\exp(0.134) - 1 = 0.143$; at WE $= 2$: $\exp(0.090) - 1 = 0.094$.

DRAFT