



**HAL**  
open science

## Trends in Social Mobility in Post-Revolution China

Yu Xie, Hao Dong, Xiang Zhou, Xi Song

► **To cite this version:**

Yu Xie, Hao Dong, Xiang Zhou, Xi Song. Trends in Social Mobility in Post-Revolution China. 2021.  
hal-03403420

**HAL Id: hal-03403420**

**<https://sciencespo.hal.science/hal-03403420v1>**

Preprint submitted on 26 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

LIEPP Working Paper  
October 2021, n°126

## Trends in Social Mobility in Post-Revolution China

**Yu XIE**

Peking University and Princeton University  
[yuxie@princeton.edu](mailto:yuxie@princeton.edu)

**Hao DONG**

Peking University  
[dongh@pku.edu.cn](mailto:dongh@pku.edu.cn)

**Xiang ZHOU**

Harvard University  
[xiang\\_zhou@fas.harvard.edu](mailto:xiang_zhou@fas.harvard.edu)

**Xi SONG**

University of Pennsylvania  
[xisong@sas.upenn.edu](mailto:xisong@sas.upenn.edu)



Distributed under a Creative Commons Paternité.  
Partage selon les Conditions Initiales 4.0 International License

[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)

How to cite this publication:

XIE, Yu, Hao DONG, Xiang ZHOU and Xi SONG, **Trends in Social Mobility in Post-Revolution China**, *Sciences Po LIEPP Working Paper* n°126, 2021-10-20.

# Trends in Social Mobility in Post-Revolution China<sup>1</sup>

**Yu Xie**

*Peking University and Princeton University*

**Hao Dong**

*Peking University*

**Xiang Zhou**

*Harvard University*

**Xi Song**

*University of Pennsylvania*

## **Abstract**

*In this paper, we study long-term trends in social mobility in the People's Republic of China since its inception in 1949, with 2 operationalizations: (1) intergenerational occupational mobility, and (2) intergenerational educational mobility. We draw on an accumulation of administrative and survey data and provide comparable estimates of these measures for birth cohorts born after 1945. To help interpret the results, we compare trends in China to those in the US for the same birth cohorts. We find an increase in intergenerational occupational mobility in China due to its rapid industrialization in recent decades. Net of industrialization, however, intergenerational occupational mobility has been declining for recent cohorts. Intergenerational educational mobility in China shows a similar declining trend. In addition, mobility patterns have differed greatly by gender, with women in earlier cohorts and from a rural origin particularly disadvantaged. We attribute the general decline in social mobility to market forces that have taken hold since China's economic reform that began in 1978. In contrast, social mobility by both measures has been relatively stable in the US. However, while social mobility in China has trended downward, it is still higher than that in the US, except for women's educational mobility.*

**Keywords:** Social mobility, trends, occupation, education, China, United States.

---

<sup>1</sup> Direct all correspondence to Yu Xie, Department of Sociology, Princeton University, 186 Wallace Hall, Department of Sociology, Princeton University, Princeton, NJ 08544, USA (email: [yuxie@princeton.edu](mailto:yuxie@princeton.edu)) or Hao Dong, Center for Social Research, Peking University, 5 Yiheyuan Road, Haidian, Beijing 100871, China (email: [dongh@pku.edu.cn](mailto:dongh@pku.edu.cn)).

The research was partially supported by the Paul and Marcia Wythes Center on Contemporary China at Princeton University and the Guanghua School of Management at Peking University. Yu Xie's work was also supported by the Laboratory for Interdisciplinary Evaluation of Public Policies (LIEPP) (ANR-11-LABX-0091, ANR-11-IDEX-0005-02 et ANR-18-IDEX-0001) at Sciences Po, where he was a visiting scholar in 2019. The ideas expressed herein are those of the authors.

## Introduction

Intergenerational social mobility, or simply social mobility, refers to the extent to which the social status of individuals, i.e., social destination, resembles that of their parents, i.e., social origin. Social mobility is a pressing issue of great concern to both academics and policy makers in an era of rising inequality around the globe. It takes on particular importance for contemporary China because the Chinese Communist Party, the ruling party since the founding of the People's Republic of China (PRC) in 1949, has explicitly promoted social mobility as part of its core communist ideology (1). In 2019, Xi Jinping's government launched a propaganda campaign promoting the benefits of high social mobility, chief among which are economic efficiency in the allocation of talent and political stability (2); this was followed by new government directives to increase social mobility (3).

What have been the historical trends in social mobility in China since the founding of the PRC in 1949? In this paper, we provide systematic evidence to answer this question, focusing on intergenerational occupational and educational mobility for cohorts born after 1945. To aid interpretation, we compare trends in China to those in the US using the same metrics and for the same cohorts. It has now been well documented that income inequality in China rose rapidly during the 1990s and 2000s, to a level much higher than that of the US (4). The "Great Gatsby Curve" hypothesis, based on cross-sectional data, suggests a negative association between income inequality and social mobility (5, 6), leading us to expect social mobility to have declined in China to a level lower than the US.

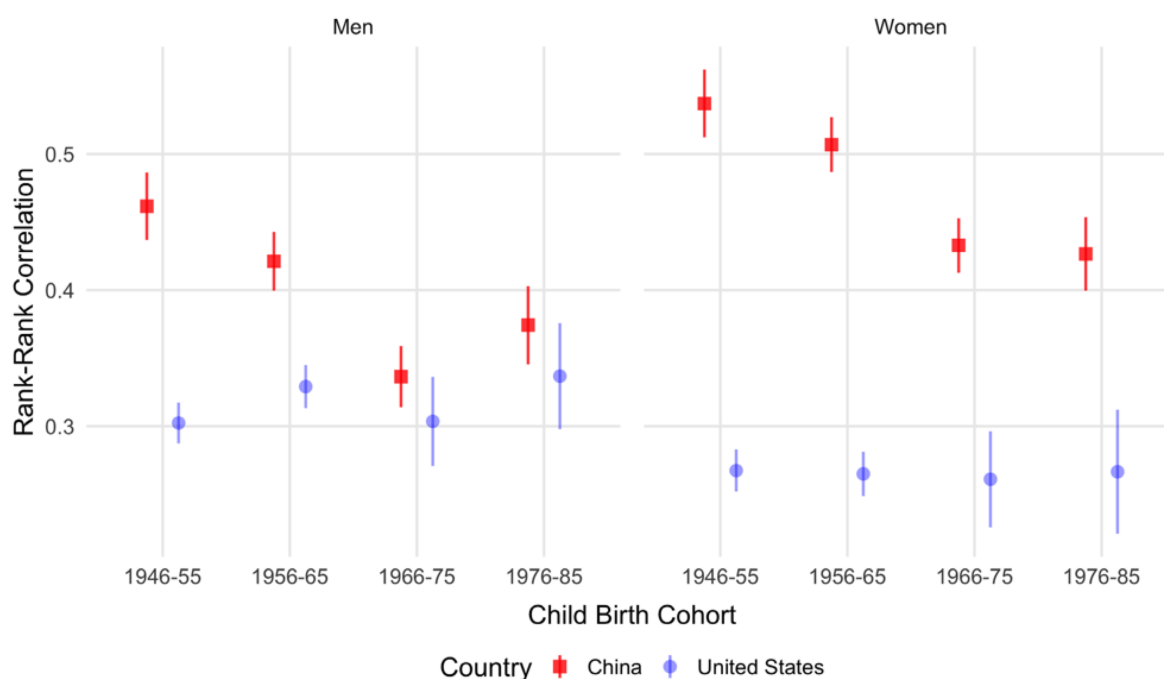
However, longitudinal evidence reveals that social inequality and social mobility are 2 different dimensions of social stratification that do not always exhibit concomitant patterns, as hypothesized in the "Great Gatsby Curve." In the US, income inequality has exploded since the 1980s, but intergenerational mobility has remained largely stable. For example, between 1980 and 2014, the pre-tax income of the bottom 50% of American adults grew by only 1%, in contrast to that of the top 10%, which grew by 121% (7). During the same period, social mobility, measured using the parent-offspring association in relative occupation and income ranks, has been largely unchanged (8, 9). In this study, we provide a systematic analysis of trends in social mobility for both men and women in China and the US, utilizing a multitude of data sources from administrative records and social surveys with rank-based measures of occupation and education that are comparable across birth cohorts, generations, genders, and the 2 countries. These rank-based measures of occupation and education reflect the respondents' and their parents' relative social standing among peers, unaffected by large social changes in either occupational structure or educational distribution over a long span of time.

Existing work on trends in intergenerational occupational mobility in the US has mostly focused on men (9), because a large portion of American women did not participate in the labor force prior to the 1960s. However, the situation has been very different in China, as "women's employment has been nearly universal throughout the history of the PRC" (10). For this reason, we examine long-term trends in intergenerational occupational and educational mobility for both genders.

## I. Results

### 1.1. Intergenerational Occupational Mobility

In Fig. 1, we present correlation coefficients for the association in occupational percentile ranks between fathers and offspring for all workers by gender and birth cohort. For simplicity, we call the quantity “rank–rank correlation”: A higher rank–rank correlation means a higher resemblance between father’s and child’s status—thus a lower level of intergenerational social mobility. Fig. 1 reveals a cohort trend in China that is distinct from that in the US. Confirming earlier research using either income-based (8) or occupation-based ranks (9), we find the rank–rank correlation to have been relatively stable in the US. Our correlation estimates, around 0.3, closely resemble those in an earlier study by Hout using occupational socioeconomic index (12: Fig. 1). Consistent with Hout’s study, our results show father–son correlations to be higher (above 0.3) than father–daughter correlations (below 0.3). In contrast, the rank–rank correlation started at a much higher level in China for the cohort born in 1946–55, at 0.46 for men and above 0.54 for women, and has trended downward for both genders, indicating a steady improvement in social mobility. The numerical results are given in *SI Appendix*, Table S1.

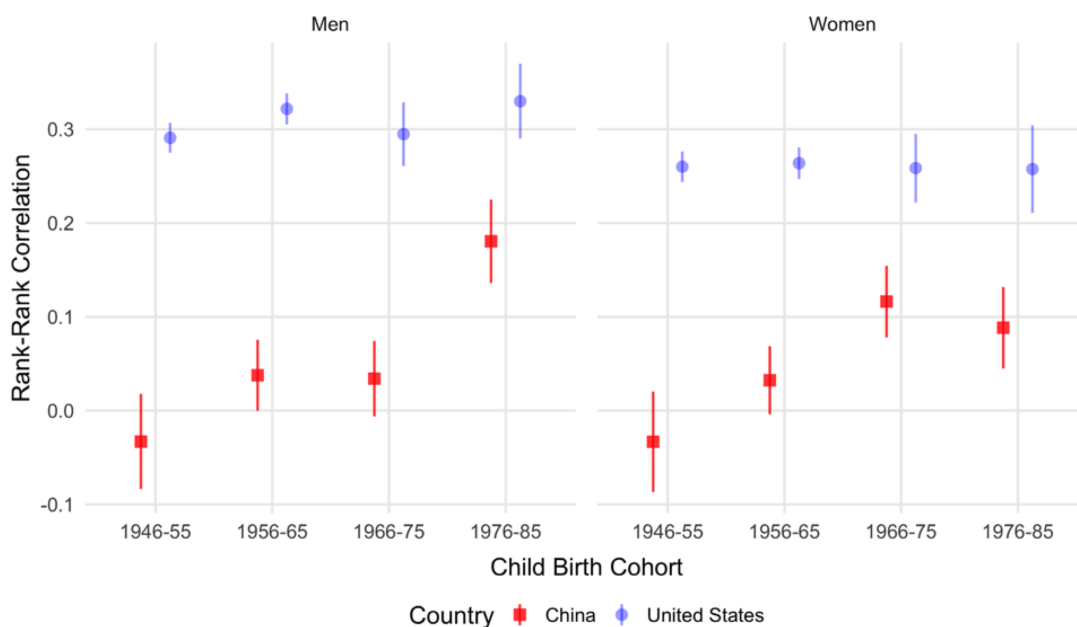


**Fig. 1.** Trends in intergenerational occupational mobility for all workers, China and US compared. *Data sources:* see *SI Appendix*, Subsections S1–S4.

It has been established that occupational inheritance is particularly strong for agricultural workers, because only children of agricultural workers are likely to be agricultural workers themselves (13). Thus, the low intergenerational occupational mobility, or high origin–destination correlation, observed for earlier cohorts in China can be attributed to low levels of industrialization prior to China’s economic reform that began in 1978. That is, China’s labor force was overwhelmingly agricultural before the economic reform and underwent a major shift away from agriculture during China’s post-reform industrialization

(14). In contrast, the US has been fully industrialized throughout the period examined, with less than 3% of the labor force working as farmers or farm laborers (13). Because land in China is owned by the state or by collectives, the distinction between farmers and farm laborers is ambiguous. For simplicity, below, we use the term “farmers” to refer to agricultural workers in general, including farmers, farm laborers, stock-raisers, animal breeders, florists, and others employed in the agricultural/horticultural sectors.

To understand China’s trends in occupational mobility net of industrialization, we reanalyzed the data after excluding workers from a farm origin, i.e., workers whose fathers were farmers. We present the new results in Fig. 2, with numerical results given in *SI Appendix*, Table S1. The results for the US are unaffected by the change. The trend for China, however, is reversed. Among children of non-farmers, the rank–rank correlation in China has trended upward from around 0 for the first cohort, born in 1946–55, to about 0.2 for men and 0.1 for women in the most recent cohort, born in 1976–85. By the American standard, shown in the same figure, these are exceptionally low correlations, indicating very weak (but growing over time) associations between father’s and child’s occupational status in China. The very low origin–destination correlations for the earlier cohorts in part reflect the rupture in the order of social stratification due to the Communist Revolution that culminated in the founding of the PRC in 1949 (1, 15). Although the overall trend is one of convergence with the US, intergenerational occupational mobility for children of non-farmers is still higher in China than in the US, even for the most recent birth cohorts. These results confirm the findings of an earlier study (14).



**Fig. 2.** Trends in intergenerational occupational mobility excluding workers from farm origin, China and US compared. *Data sources:* see *SI Appendix*, Subsections S1–S4.

Whereas the rank–rank correlation has been consistently higher for men than for women in the US, the gender gap is inverted for all Chinese workers, with the rank–rank correlation much higher for women than for men, as shown in Fig. 1. In Fig. 2, where we restricted the analysis to workers from a non-farm origin, the gender difference for China disappears. This change indicates that Chinese women from a farm origin have experienced

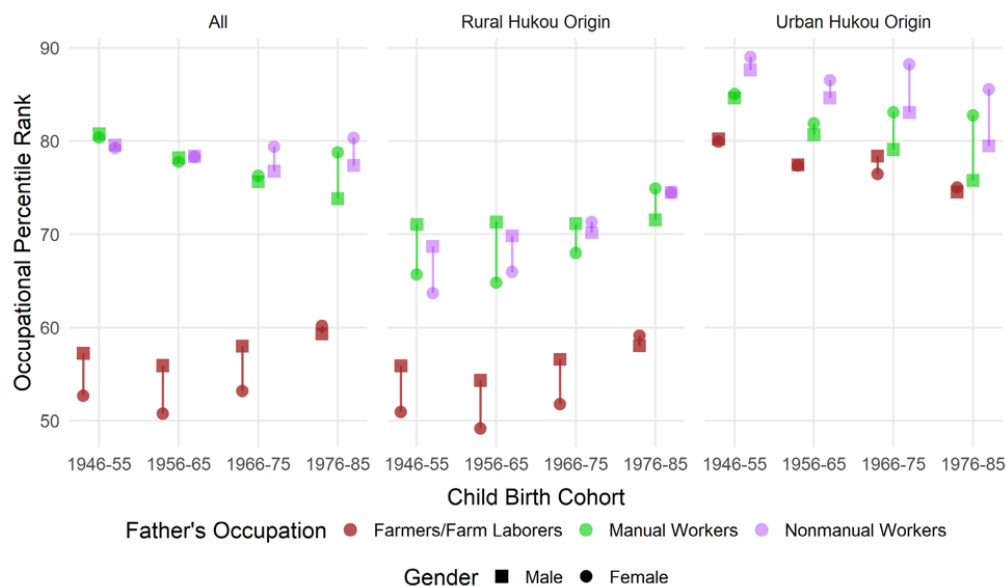
significantly lower social mobility relative to men. Given the demographic history that fertility was relatively high and sex-selective abortion was largely absent for most Chinese cohorts covered in this study (16), we know that most rural families had children of both genders, with the gender composition unrelated to parental socioeconomic status. In this context, the existence of a very large gender gap in favor of men's mobility in Fig. 1, and its absence in Fig. 2, reveal strong son preference among farming families in promoting social mobility. There was an economic rationale for this gender inequality. In the traditional Chinese patriarchal family system, sons are permanent members of their natal family and retain lifetime financial relationships with their parents. Daughters become contributors to their husband's family upon marriage. Thus, it was in parents' self-interest to invest in sons rather than daughters (10).

At first glance, the main explanation for a preference for the social mobility of sons among Chinese farming families seems to be China's patriarchal family tradition. However, this explanation overlooks the fact that the Chinese family institution has been challenged and repudiated repeatedly by several major social movements in the modern history of China: most notably the May Fourth Movement in 1919, the Communist Revolution that resulted in the founding of the PRC in 1949, the 1966–76 Cultural Revolution, and the economic reform that began in 1978 (10). As a result of these large-scale social movements, the family in contemporary China has fundamentally changed, now resembling that in developed East Asian countries such as Japan and South Korea, and to a large extent even the US and European countries (17).

We believe that a crucial reason for the continuation of son preference in terms of social mobility among post-1949 Chinese farming families is *hukou*, a government-controlled institution of social stratification that is unique to China (18). While literally translated as household registration, *hukou* is the system through which the government regulates where Chinese people live. Soon after the founding of the PRC in 1949, the *hukou* system separated the Chinese people as if they belonged to 2 castes, rural and urban, with the latter having social privileges not enjoyed by the former, such as food provision, public housing, comprehensive medical care, better schooling, job assignments, and old-age pension. For rural Chinese, upward mobility has often meant the conversion from rural *hukou* to urban *hukou* (18), say through higher education or military service. Earlier research has shown that the implementation of a state pension system for old-age support has removed the incentives for son preference in urban China (19).

We test our conjecture that *hukou* accounts for China's large gender gap in social mobility shown in Fig. 1 with a more detailed analysis. In Fig. 3, we present trends in occupational status by father's occupation (measured in 3 broad categories: farmer, manual worker, and non-manual worker), child's gender, and child's *hukou* origin (i.e., *hukou* status in childhood). Detailed numerical results are given in *SI Appendix*, Table S2. We observe that rural *hukou* origin and farm origin are both associated with lower occupational status compared to other groups. For example, for the birth cohort 1946–55, men of rural *hukou* origin whose fathers were farmers attained an average occupational percentile score of 56, compared to 88 for men of urban *hukou* origin whose fathers were non-manual workers. A striking pattern, however, pertains to the changing gender gap by *hukou* origin: for all birth

cohorts, except the most recent one, 1976–85, there was clearly a gender gap in favor of men in terms of occupational status for persons of rural *hukou* origin, across all 3 categories of father’s occupation. In contrast, we do not find a discernable disadvantage for women in terms of occupational status for those of urban *hukou* origin, regardless of father’s occupation. In fact, for the most recent 2 cohorts, i.e., those born after 1966, women of urban *hukou* origin whose fathers were either manual or non-manual workers attained a higher occupational status on average than their male counterparts, *ceteris paribus*.



**Fig. 3.** Trends in occupational percentile rank by gender and father’s occupation among all individuals, individuals of rural *hukou* origin, and individuals of urban *hukou* origin. *Data sources:* see *SI Appendix*, Subsections S1–S4.

To illustrate the particular disadvantage of women of rural *hukou* origin, we further examine specific occupational destinations by gender, *hukou* origin, and cohort. The detailed results are given in *SI Appendix*, Table S3. Again, we find clear disadvantages for women of rural *hukou* origin relative to men, but not for women of urban *hukou* origin. For the first cohort, for example, women of rural origin are much more likely to be farmers than their male peers (0.77 versus 0.66), and much less likely to become high-status white-collar workers (managers, professional workers, and large proprietors) (0.03 versus 0.09). Over successive cohorts, as the percentage of the farming population declined and the percentage of high-status white-collar jobs increased, the gender disparity narrowed. These results not only confirm our earlier finding that rural Chinese in earlier cohorts had limited social mobility overall, but also reveal particular disadvantages experienced by women of rural origin in earlier cohorts.

From the above results, we draw the conclusion that Chinese women experienced lower social mobility than men only if they were of rural *hukou* origin. This gender disparity was particularly pronounced in the earlier cohorts. With time, the institutional effect of *hukou* on gender inequality has been eroded.

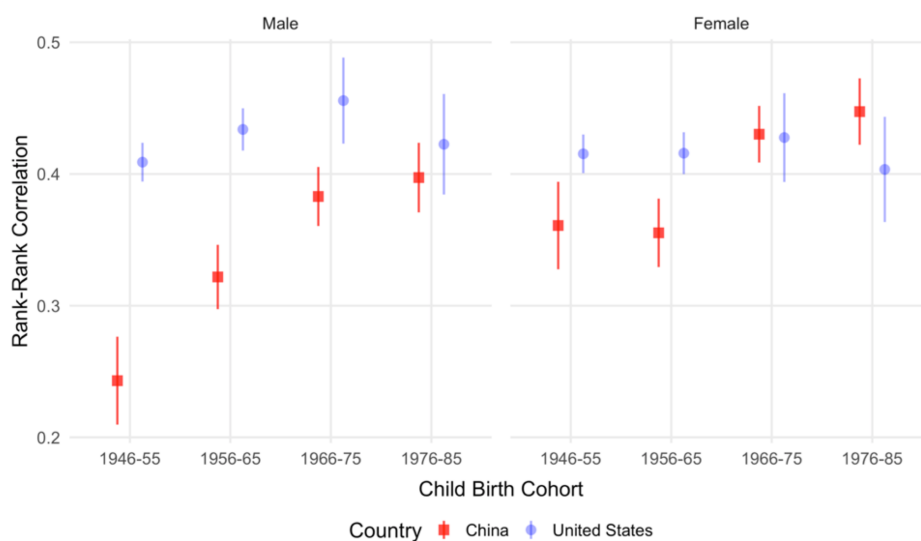
## 1.2. Intergenerational Educational Mobility

In sociological research on social mobility, occupation has been the standard choice as a measure of social status, for 2 main reasons. First, occupation is highly correlated with other



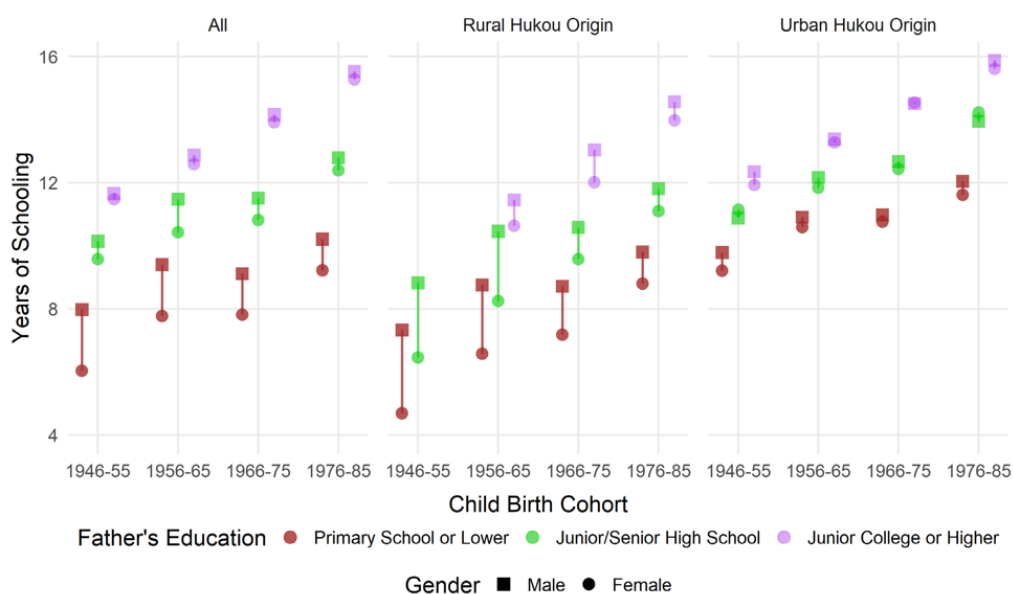
measures of socioeconomic status, such as income, wealth, and education. Second, occupation is an individual's achieved status that is typically publicly known as a primary basis for social prestige and often collected in surveys and administrative registers. However, occupation does not necessarily stay the same over the life course, as assumed by our methodology using the rank–rank correlation. Using occupation as a measure of social status also requires us to ignore potential individual-level heterogeneity within a large occupation. For example, farming accounted for more than half of the labor force in pre-reform China, but not all farmers had the same social status. In addition, occupation is not defined for persons who were unemployed or out of the labor force at the time of data collection. For these reasons, we supplement the above analyses of trends in intergenerational occupational mobility with parallel analyses of educational mobility.

We present our main findings on educational mobility in Fig. 4, with numerical results given in *SI Appendix*, Table S4. Overall, the rank–rank correlation in education is higher than that in occupation. For example, for US men, the educational correlation has been around 0.4, in contrast to an occupational correlation of around 0.3. Comparing Fig. 4 to Fig. 2, we observe a striking similarity in trends, especially for men, between intergenerational educational mobility and intergenerational occupational mobility in the non-farm origin population. The rank–rank correlation in men's educational status in China started at a much lower level (0.24) for the first cohort, 1946–55, than that in the US, and approached the level of the US (0.42) in the most recent cohort, 1976–85, at 0.40. While the change across cohorts is less pronounced, women's intergenerational educational mobility in China also appears to converge with that of their counterparts in the US. Specifically, the country-level difference in the father–daughter rank–rank correlation started small in the first cohort, at 0.36 in China versus 0.42 in the US; became equal in the third cohort, 1966–75, at 0.43; and reversed for the last cohort, at 0.45 in China and 0.40 in the US. Again, we observe lower correlations for men than for women in China using education-based ranks, echoing the results for occupation-based ranks for all workers shown in Fig. 1.



**Fig. 4.** Trends in intergenerational educational mobility, China and US compared. *Data sources:* see *SI Appendix*, Subsections S1–S4.

Similar to the earlier analysis of occupational mobility, we suggest *hukou* as an explanation for women’s lower educational mobility when compared to men. We performed a detailed analysis of the Chinese data, with results presented in Fig. 5 and numerical results given in *SI Appendix*, Table S5. In Fig. 5, we measure child’s education by average years of schooling, broken down by father’s education (measured by 3 main levels: primary school or lower, junior/senior middle school, and junior college or higher), in addition to child’s birth cohort, gender, and *hukou* origin. Of course, the most visible trend in Fig. 5 is an overall increase in the years of schooling due to a rapid expansion in education provision in China over the course of this period.



**Fig. 5.** Trends in average years of schooling by gender and father’s education among all individuals, individuals of rural *hukou* origin, and individuals of urban *hukou* origin. *Data sources:* see *SI Appendix*, Subsections S1–S4.

Note: Due to insufficient data, the average years of schooling are not computed for individuals in the earliest birth cohort who had a rural *hukou* origin and a father of junior college or higher education.

For both urban and rural *hukou* origin groups, we observe a strong gradient by father’s education, i.e., children of more educated fathers attain more years of education. There has been a persistent rural–urban gap in education. For example, for the first cohort, men of rural origin whose fathers had junior or senior middle school education attained, on average, 8.8 years of schooling, compared to 10.9 years among men of urban origin with similarly educated fathers. Across all the cohorts and for all the levels of father’s education, we do not see a noticeable gender gap for persons of urban origin. A substantial gender disparity, however, existed for all cohorts and for all 3 levels of father’s education for rural-origin individuals. Specifically, the gender disparity among Chinese of rural origin was strikingly large for the cohort born in 1946–55 but narrowed gradually over successive cohorts. Again, similar to the results for occupational mobility, we found the strong (but declining) role of *hukou* in limiting women’s educational mobility in China.

## II. Conclusion

With few exceptions (20), a vast literature has established that relative intergenerational mobility has been stable, or trendless, in industrialized societies in the West (8, 9, 13, 21). Given China's rapid rise in income inequality in the post-reform era (4), there have been concerns that intergenerational mobility has trended downward (14, 22). The topic of social mobility is also highly political in today's China, as the Chinese Communist Party bases its ruling legitimacy on the promise of delivering social openness, fairness, and mobility (1, 2). It is clear that one major theme in Xi Jinping's leadership is "common prosperity," which means reduction of inequality. Between January to August 2021, Xi made 13 public speeches where the theme of common prosperity was emphasized (23). The latest concrete measure taken by the government is the prohibition of after-school tutoring in academic subjects to precollege students (24). Is the track record in China's recent past so bad that it justifies the government's ongoing strong intervention to promote social mobility?

Drawing on an accumulation of massive administrative and survey data and using comparable measures, we have carefully examined the long-term trends of intergenerational mobility in China for birth cohorts born after 1945, i.e., those who grew up after the founding of the PRC in 1949. To help interpret the results, we compare trends in China to those in the US for the same birth cohorts. We develop 2 relative, comparable measures of social mobility: rank-rank correlations measuring intergenerational occupational mobility and intergenerational educational mobility.

Our research yields mixed results. Due to rapid industrialization, intergenerational occupational mobility in China has greatly improved over time. Net of industrialization, however, intergenerational occupational mobility has been declining for recent cohorts. If we use the education-based measure, we observe a similar decline in intergenerational educational mobility in China. While these findings lend support to the Chinese government's concern that relative social mobility has declined in China, we should note that social mobility among children of non-farm origin was exceptionally high for the earliest Chinese cohorts in our study. This is most apparent when we compare the trends in China to those in the US, which have been relatively stable. Although social mobility in China has declined for more recent cohorts, it remained higher than that in the US for the last cohort, except for women's educational mobility.

We also found large gender differences in social mobility trends for Chinese of rural *hukou* origin. In earlier cohorts, while social mobility was high for urban residents of both genders and rural men, mobility was relatively limited for rural women. Girls born in rural China were severely disadvantaged relative to their male counterparts, having much lower likelihoods of obtaining schooling, leaving farming, and entering high-status white-collar jobs. Over time, forces of industrialization, education expansion, and fertility reduction have eroded the strong limiting factor of rural *hukou* on women's social mobility, substantially narrowing, or in some aspects eliminating, the gender disparity among persons of rural *hukou* origin. While gender inequality has not fully disappeared, China no longer stands out as an outlier as it was before in limiting women's social advancement.

### III. Materials and Methods

#### 3.1. Data

To study trends in social mobility in China, we analyze data from 2 major sources: (a) the 1982, 1990, 2000, and 2010 China Censuses, and the One-Percent Population Survey of China in 2005 (henceforth the 2005 China Mini-Census) (see *SI Appendix*, Subsection S1); (b) 2 series of large-scale social surveys from the project of Life Histories and Social Change in Contemporary China in 1996 (LHSCCC1996) and the Chinese General Social Survey in 2005, 2006, 2008, 2010, 2011, 2012, 2013, 2015, 2017, 2018 (CGSS2005–2018). The LHSCCC1996 and CGSS2005–2018 data contain detailed information about respondents' own education and occupation and those of their parents when the respondents were children (see *SI Appendix*, Subsection S2). We pool the survey data to track changes over birth cohorts.

Data for the US also contain 2 major sources: (a) cross-sectional US Population Census data from 1900 to 2000 and the American Community Survey (ACS) data from 2001 to 2015 (see *SI Appendix*, Subsection S3); (b) 8 large-scale US social surveys including the General Social Survey 1972–2016; National Longitudinal Survey–Young Men 1966–1981; National Longitudinal Survey–Young Women 1968–1993; National Longitudinal Survey of Youth 79 1979–2012; National Survey of Families and Households 1987, 1993, and 2002; Occupational Changes in a Generation II (in 1973); Panel Study of Income Dynamics (SRC sample) 1968–2015; and Survey of Income and Program Participation 1986, 1987, and 1988 (see *SI Appendix*, Subsection S4). We pool the survey data that contain detailed information about each respondent's and his/her father's occupation and education, to form the main analytical sample.

Given inconsistency in occupational classification across surveys, we harmonize occupational variables from different data sources into a 2-digit occupational classification. For China, we use the occupation classification system of the 2000 China Census, which contains 71 unique occupational categories (see *SI Appendix*, Subsection S5). For the US, we harmonize occupational variables from different data sources into standard 1950 Census Bureau occupation codes (see *SI Appendix*, Subsection S6). New occupations that did not emerge until recently, such as computer programmer, computer systems analyst, and software engineer, were coded into a broader category “professional, technical & kindred workers (nec).” The standard 1950 occupational classification scheme consists of 283 occupational categories, but some of these occupations are not consistently recorded across census years. We thus map the 1950 occupations into Weeden and Grusky's microclass occupational scheme that is widely used in comparative studies on intergenerational mobility (9). The revised scheme includes 70 unique occupational categories, a number similar to that of the occupational variable used for the Chinese data.

### 3.2. Measuring Occupational Status

We measure occupational status for both social origin and destination relatively, rather than absolutely, so that our measurement is not confounded by large social changes in occupational structure over time or large structural differences in occupational structure between China and the US. With a harmonized occupational classification for each country, we convert occupation into a relative status measure based on an education-based occupational ranking, relying on 2 assumptions. First, different occupations can be rank-ordered in terms of socioeconomic status. Second, an individual's occupational status is largely stable over the life course such that the estimates of intergenerational occupational mobility do not depend on the age at which occupation is measured. Although contestable, these assumptions are good approximations to social reality and have been widely used in prior research (9).

For each country, we pool all available census data (including the 2005 China Mini-Census for China and the ACS samples for the US) and generate occupational percentiles in 4 steps. First, we partition the data into 10-year birth cohorts. Within each of these 10-year cohorts, we rank individuals according to their level of education and calculate the average percentile rank for each occupational group, which can be viewed as an occupational status score. Next, we rank all occupations within each cohort based on these status scores, accounting for relative occupational sizes. The resulting occupational percentile rank (0–100) represents a person's relative socioeconomic status within a birth cohort. A higher percentile rank indicates a higher socioeconomic status. To assess intergenerational mobility of occupational status, we calculate occupational percentile ranks for all parents and children in the survey data, according to their respective birth cohorts. See *SI Appendix*, Subsection S7 for more methodological details.

### 3.3. Measuring Educational Status

To study intergenerational educational mobility, we similarly derive relative measures of education for both social origin and destination so as to achieve measurement comparability over cohorts, between parents and children, between genders, and between China and the US. Normalization within gender is necessary given women's rapid progress in educational achievement in China (11). We convert observed educational attainment to percentile ranks within cohort and gender for each country. The number of educational categories is fewer than that of occupation. We construct 7 levels of educational attainment for China and 11 for the US. Since these levels of education are already ranked, there is no need to derive a rank order.

We calculate educational percentile ranks in 3 steps. First, we pool available administrative data (Census/ACS for the US and Census/Mini-Census for China) for each country and partition the data into 10-year birth cohorts described above. Second, within each country and census wave, we rank men and women separately by their highest grade completed to form cumulative education distributions. We use the midpoint to adjust for percentile ranks of individuals with the same level of educational attainment. The resulting percentile rank represents the relative status of an individual by his or her educational attainment within the same-gender, 10-year birth cohort. To examine educational mobility, we assign these percentile ranks to individuals and their parents in the corresponding analytical

samples, according to their country, birth cohort, gender, and educational attainment. See *SI Appendix*, Subsection S8 for more methodological details.

## Bibliography

- CHETTY, R., *et al.*, Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *Am. Econ. Rev.* 104, 141–47 (2014).
- CORAK, M., The economics of the Great Gatsby Curve. Personal blog, 17 January 2012. Available at <http://mileskorak.com/2012/01/17/the-economics-of-the-great-gatsby-curve/>. Accessed 4 August 2021.
- FAN, Y., & YI, J., & ZHANG, J., Rising intergenerational income persistence in China. *Am. Econ. J.: Econ. Policy* 13, 202–230 (2021).
- GANZABOOM, H. B. G., & TREIMAN, D. J., & UTLEE, W. C., Comparative intergenerational stratification research: Three generations and beyond. *Ann. Rev. Sociol.* 17, 277–302 (1991).
- HOUT, M., Americans' occupational status reflects the status of both of their parents." *Proc. Natl. Acad. Sci. U.S.A.* 115, 9527–9532 (2018).
- JINPING, X., *Gongtong fuyu shi shehuizhuyi de benzhi yaoqiu* [Common prosperity is the essential requirement of socialism]. Qizhiwang.org.cn, 25 August 2021. Available at <http://www.qizhiwang.org.cn/n1/2021/0825/c422351-32207065.html>. Accessed 5 September 2021.
- KRUEGER, A. B., The rise and consequences of inequality in the United States. Prepared remarks, 12 January 2021. Available at [https://obamawhitehouse.archives.gov/sites/default/files/krueger\\_cap\\_speech\\_final\\_remarks.pdf](https://obamawhitehouse.archives.gov/sites/default/files/krueger_cap_speech_final_remarks.pdf). Accessed 4 August 2021.
- LI, Y., The important code to contemporary China's prosperity and development. *People's Daily*, 26 July 2019. Available at <http://theory.people.com.cn/n1/2019/0726/c40531-31257113.html>. Accessed 4 August 2021.
- Ministry of Human Resources and Social Security of PRC, Promoting labor force and talent social mobility and stimulating latent forces for economic and social development. 10 January 2020. Available at [http://www.mohrss.gov.cn/SYrlzyhshbzb/zcfg/SYzhengcejiedu/202001/t20200110\\_353697.html](http://www.mohrss.gov.cn/SYrlzyhshbzb/zcfg/SYzhengcejiedu/202001/t20200110_353697.html). Accessed 4 August 2021.
- PENG, X. China's demographic history and future challenges. *Science* 333, 581–587 (2011).
- PIKETTY, T., & SAEZ, E., & ZUCMAN, G., Distributional national accounts: Methods and estimates for the United States. *Q. J. Econ.* 133, 553–609 (2018).
- SONG, X., *et al.*, Long-term decline in intergenerational mobility in the United States since the 1850s. *Proc. Natl. Acad. Sci. U.S.A.* 117, 251–258 (2020).
- STEVENSON, A. & LI, C., China targets costly tutoring classes. Parents want to save them. *New York Times*, 30 July 2021.
- TREIMAN, D. J., Trends in educational attainment in China. *Chinese Sociol. Rev.* 45, 3–25 (2013).
- TREIMAN, D. J., & WALDER, A. G., The impact of class labels on life chances in China. *Am. J. Sociol.* 124, 1125–1163 (2019).
- VALLET, L-A., Forty years of social mobility in France. *Rev. française sociol.* 42, 5–64 (2001).
- WU, X., & TREIMAN, D. J., The household registration system and social stratification in China: 1955–1996. *Demography* 41, 363–384 (2004).
- XIE, Y., "Gender and family" in *The Oxford companion to the economics of China*, S. Fan, R. Kanbur, S.-J. Wei, X. Zhang, Eds. (Oxford: Oxford University Press, 2014), 495–501.

XIE, Y., & KILLEWALD, A. A., Intergenerational occupational mobility in Britain and the U.S. since 1850: Comment. *Am. Econ. Rev.* 103, 2003–2020 (2013).

XIE, Y., & ZHANG, C. The long-term impact of the Communist Revolution on social stratification in contemporary China. *Proc. Natl. Acad. Sci. U.S.A.* 116, 19392–19397 (2019).

XIE, Y., & ZHOU, X., Income inequality in today's China. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6928–6933 (2014).

XIE, Y., & ZHU, H., Do sons or daughters give more money to parents in urban China? *J. Marriage Fam.* 71, 174–186 (2009).

YU, J., & XIE, Y., Recent trends in the Chinese family: National estimates from 1990 to 2010. *Demogr. Res.* 44, 595–608 (2021).

ZHOU, X. & XIE, Y., Market transition, industrialization, and social mobility trends in post-revolution China.” *Am. J. Sociol.* 124, 1810–1847 (2019).



## **Appendix**

Supplementary Information (SI) for

### **Trends in Social Mobility in Post-Revolution China**

**This PDF file includes:**

Materials and Methods Sections S1 to S8

Tables S1 to S5

References

## **S1 China Census Data**

To create cohort-specific occupational percentile ranks, we used individual-level data from the 1982, 1990, and 2000 China Censuses in the Integrated Public Use Microdata Series (IPUMS hereafter) and the One-Percent Population Survey of China in 2005 (also called the “2005 China Mini-Census”). The IPUMS project provides harmonized China Census microdata, in which all census records have been converted into a format with consistent variable names, sample restrictions, coding methods, and documentation. We pooled individuals who were born in the same year across different census years. We first restricted the sample to males and females aged 25 to 64 and then generated occupational percentile ranks by birth cohort based on occupation-specific educational distributions.

Similarly, we used these 3 waves of census microdata plus aggregate-level tabulations from the 2010 China Census to create educational percentile ranks specific to each cohort and gender. We used the IPUMS-harmonized variable in the microdata samples, EDUCCN, measuring the highest educational level that a person had attained. We restricted each sample to males and females aged 25 and above. When using tabulation data from the 2010 Census, we calculated the educational percentile ranks in the same fashion. For each cohort covered by multiple census waves, we chose the educational distribution from the census wave in which the cohort would be closest to the age range 30–40, so that most individuals in the cohort would have completed their education while potential risk for mortality selection by education is at the minimum. We generated educational percentile ranks for men and women separately.

## **S2 Large-Scale Social Survey Data in China**

### **Life Histories and Social Change in Contemporary China**

The project Life Histories and Social Change in Contemporary China was a nationally representative survey conducted by UCLA, Stanford University, and Hong Kong University of Science and Technology in collaboration with Renmin University of China in 1996 (hereafter LHSCCC1996). The survey was designed to collect data from a national probability sample of Chinese adults to permit comparisons across cohorts in domains such as education, income, occupation, housing and property, and life histories. Urban and rural areas were sampled separately, and interviews were completed for 3,087 urban residents and 3,003 rural residents aged 20–69 in 1996. We applied the sample weight (*weight*) to account for the multistage sampling design and restricted our analysis to only respondents aged 25 to 64.

In LHSCCC1996, the respondent was asked to report his/her own occupation at the time of the survey as well as his/her father’s occupation when the respondent was at age 14. The final analysis for occupational mobility included 1,436 male respondents and 1,221 female respondents. Information was also solicited for the respondent’s own and his/her father’s highest level of education. The final analysis for educational mobility included 2,568 male respondents and 2,472 female respondents.

The data are publicly available and can be downloaded from the LHSCCC website:  
<https://www.library.ucla.edu/social-science-data-archives/life-histories-social-change-china>.

### Chinese General Social Survey

The Chinese General Social Survey (CGSS) is a nationally representative, repeated, cross-sectional survey that has been conducted by the National Survey Research Center at Renmin University since 2003. The CGSS survey is designed to monitor changes in social structure and quality of life of urban and rural households in China annually or biennially. CGSS uses 3 different probability proportional to size (PPS) sampling schemes: the 2003–2006 sampling scheme, the 2008 experimental sampling scheme, and the 2010–2019 sampling scheme. In the 2003–2006 sampling scheme, 10 households were randomly selected from each of 1,000 communities, with 2 communities selected from each of 500 enumeration districts, with 4 enumeration districts selected within each of 125 primary sampling units. In 2008, 10 households were randomly selected from each of 600 communities, with 2 communities selected within each of 300 enumeration districts, with 3 enumeration districts selected within each of 100 primary sampling units. In the 2010–2019 sampling scheme, 25 households were randomly selected from each of 480 enumeration districts, with 2–4 enumeration districts selected within each of the 140 primary sampling units. A respondent aged 18–69 was randomly selected from each household to participate in a personal interview. We used data from the years 2005, 2006, 2008, 2010, 2012, 2013, 2015, 2017, and 2018 and applied the sample weight (*weight*) to account for the multistage sampling design.

In each survey year, the respondent was asked to report his/her current occupation as well as his/her father's occupation when the respondent was at age 14 (or age 18 in CGSS 2006). The final analysis for occupational mobility included 21,569 male respondents and 22,155 female respondents from CGSS 2005–2018. Individuals were also asked to report their own highest levels of education and those of their fathers. The final analysis for educational mobility included 28,098 male respondents and 29,211 female respondents from CGSS 2005–2018.

The data are publicly available and can be downloaded from the CGSS website:  
<http://cnsda.ruc.edu.cn/index.php?r=site/datarecommendation>.

### S3 US Census Data

We first used the US Population Census data from 1900 to 2000 and American Community Survey (ACS) data from 2001 to 2016 in IPUMS to create occupational percentile ranks by birth cohort, as in a prior study by Song *et al.* (1). IPUMS USA provides harmonized US Census microdata, in which all census records have been converted into a format with consistent variable names, sample restrictions, coding methods, and documentation. For years for which the full-count census data were available (e.g., 1940), we chose the full-count data over census samples. For years for which both 1% and 5% samples were available (e.g., 2000), we chose the larger sample. For the year 1970, 6 1% samples were drawn independently from the population data. We included the 2 1% samples known as Form 1 and Form 2 Metro in our analysis. We pooled individuals who were born in the same year but observed in different census years. We first restricted the sample to males and females aged 25 to 64 and then

generated occupational percentile ranks by birth cohort based on the literacy rate and educational distribution within an occupation.

Similarly, we used the IPUMS Population Census data from 1940 to 2000 and ACS data from 2001 to 2016 to create educational percentile ranks in the population. Data of census waves before 1940 were not used because they included no information on detailed educational attainment comparable to that in later waves. We used the general version of the IPUMS-harmonized variable, EDUC, for an individual's educational attainment, measured by the highest year of schooling completed. For this calculation of educational percentile ranks, we used the 1% samples of the 1940 and 1970 (Form 2 Metro) censuses, and 5% samples of the 1960 and 1980–2000 Censuses. We restricted each sample to males and females aged 25 and above. We took the average for each birth cohort, gender, and educational level when corresponding percentile ranks were calculated from data of multiple available census waves.

## **S4 Large-Scale Social Survey Data in the US**

Our US analyses included 8 large-scale social surveys that have been extensively used in previous research on intergenerational social mobility. We weighted each sample using the sample weight variable included in the original data and created a cross-sample weight variable to adjust for variations in sample size across datasets.

### **General Social Survey**

The General Social Survey (GSS) is a large-scale, cross-sectional survey that has been implemented since 1972 by the National Opinion Research Center (NORC) at the University of Chicago. The survey was conducted annually in most years before 1994 and changed to a biennial basis thereafter. It is designed to be representative of all adults in US households aged 18 and older. The data have been widely used in studies on societal changes in individual attributes and attitudes and population composition. The present analysis relies on data from the years 1972 to 2018 and applies sample weight (*wtssall*) to account for the multistage sampling design. Oversamples of black respondents in GSS 1982 and 1987 are dropped from our analyses.

Respondents were asked to report their own occupations at the time of the survey and their fathers' occupations while they "were growing up." To generate individuals' occupational percentile ranks, we first derived respondents' birth cohorts using their ages in the survey year and converted their current occupations to standardized 1950 occupation codes. Next, we mapped the 1950 occupations to the microclass scheme using the same method as in the study by Song *et al.* (1). Finally, we merged the GSS data with the occupational percentile rank file generated from the population census data based on microclass occupations and birth cohort. The variable of father's birth year contains many missing values because the question was asked only in years 1993 to 1998. To impute missing data, we assumed that the father's age at the child's birth follows a truncated normal distribution varying from ages 15 to 60, with the mean equal to 30 and the variance equal to 6.9. We then subtracted father's year of becoming a parent from respondent's birth year to obtain father's birth year. The final analysis of occupational mobility included 16,608 cases

with valid data from male respondents and 18,886 cases from female respondents aged 25 to 64 and their reported fathers for the years 1972–2018.

Respondents were asked to report their own education and their fathers' education while they "were growing up." To generate individuals' educational percentile ranks, we merged the GSS data with the educational percentile rank file generated from the population census data based on educational attainment, gender, and birth cohort. The final analysis of educational mobility included 15,187 males and 17,958 females aged 25–64 from years 1972–2018.

The data are publicly available and can be downloaded from the GSS website: <http://gss.norc.org/get-the-data/>.

### **National Longitudinal Survey–Young Men**

The National Longitudinal Survey–Young Men (NLS–YM) was another dataset among the several longitudinal cohort studies conducted by the Bureau of Labor Statistics under the United States Department of Labor between 1966 and 1981. The project began with a nationally representative sample of 5,225 American males aged 14 to 24 in 1966 and was discontinued in 1981 when the respondents were 29 to 39, at the time of their last interview. Respondents were surveyed annually between 1966 and 1971, and in the subsequent years of 1973, 1975, 1976, 1978, 1980, and 1981. The data have been used to study education and labor market experiences of men during their early careers, such as their educational attainment and expectations, school-to-work transitions, labor market attachment, and activities related to crime, delinquency, and school discipline. We used data from the years 1966 to 1981 (12 rounds) and applied the sample weight (*R0000200*) to account for the multistage sampling design. We kept male respondents aged 25 and 65 in every round of the survey.

Respondents were asked to report their current or last occupations and those of their fathers in each survey year between 1966 and 1969. To make the data comparable with other datasets used in the analysis, we relied on respondents' occupations reported in the last wave of the survey and their fathers' occupations reported in 1966. If father's information was missing in the first wave, occupation in the next available year was used. We merged the NLS–YM data with the occupational percentile rank file generated from the population census data using microclass occupations and birth cohort. Because father's birth year was not asked, we impute missing data for father's birth year using the same imputation method described in the GSS data section above. The final analysis of occupational mobility included 3,506 cases with valid data from male respondents and their reported fathers from 1966 to 1981.

Respondents were asked to report their highest levels of education in each survey year and those of their fathers in 1966. We relied on respondents' highest education ever reported across survey years. We merged the NLS–YM data with the educational percentile rank file generated from the population census data using educational attainment and birth cohort. The final analysis of educational mobility included 3,927 cases with valid data from male respondents and their reported fathers from 1966 to 1981.

The data are publicly available and can be downloaded from the NLS website: <http://www.nlsinfo.org/investigator/pages/login.jsp>.

### **National Longitudinal Survey–Young Women**

The National Longitudinal Survey–Young Women (NLS–YW) was one of the several longitudinal cohort studies conducted by the Bureau of Labor Statistics under the United States Department of Labor between 1968 and 1993. The project began with a nationally representative sample of 5,159 American women aged 14 to 24 in 1968. Respondents were surveyed annually between 1968 and 1973, and in the subsequent years of 1975, 1977, 1978, 1980, 1982, 1983, 1985, 1987, 1988, 1991, and 1993. The data have been used to study education and labor market experiences of women during their early careers, such as their educational attainment and expectations, school-to-work transitions, labor market attachment, and attitudes toward working, housekeeping, and childcare. We used data from the years 1968 to 1993 (17 rounds) and applied the sample weight (*R0000200*) to account for the multistage sampling design. We kept female respondents aged 25 to 65 in every round of the survey.

Respondents were asked to report their current or last occupations in each survey year between 1968 and 1993. They reported their fathers' occupations in 1968. To make the data comparable with other datasets used in the analysis, we relied on respondents' occupations reported in the last wave of the survey. If the respondent's information was missing in the first wave, occupation in the next available year was used. We merged the NLS–YW data with the occupational percentile rank file generated from the population census data using microclass occupations by birth cohort. Because father's birth year was not asked, we impute missing data for father's birth year using the same imputation method described in the GSS data section. The final analysis of occupational mobility included 3,310 cases with valid data from female respondents and their fathers from 1968 to 1993.

Individuals were asked to report their current or last levels of education in each survey year and their fathers' highest levels of education in 1968 and 1978. We chose respondents' and their fathers' highest education across waves. The final analysis of educational mobility included 3,984 cases with valid data from female respondents and their fathers from 1968 to 1993.

The data are publicly available and can be downloaded from the NLS website: <http://www.nlsinfo.org/investigator/pages/login.jsp>.

### **National Longitudinal Survey of Youth 1979**

The National Longitudinal Survey of Youth 1979 (NLSY79) was another longitudinal cohort study conducted by the Bureau of Labor Statistics under the United States Department of Labor. The project followed the lives of a nationally representative sample of 12,686 respondents born between 1957 and 1964. The respondents were aged 14–22 when first interviewed in 1979 and revisited annually until 1994 and biennially thereafter. We used data from 1979 to 2012 and applied sample weight (*R0216100*) to account for the multistage

sampling design. We restricted our analysis to respondents aged 25 to 64 in every round of the survey.

Respondents were asked to report their occupations in each survey year and their fathers' (or stepfathers') occupations in 1978 during the first wave. In order to generate occupational percentile ranks, we calculated individuals' birth years using age records in every round of the survey and converted their occupations reported at the age closest to 40 among all the waves to standardized 1950 occupation codes and then to microclass occupations. Father's birth year was directly observed if they coresided with their child in 1979 or was indirectly reported by their child if the father was still alive in 1987 or 1988. Because the age records were reported multiple times, we converted ages into birth years and calculated the modes of sons' and fathers' birth years. We imputed missing data for father's birth year using the same imputation method described in the GSS data section. The final analysis for occupational mobility relied on 4,136 cases with valid data on male respondents and 4,038 female respondents and their fathers.

Individuals were asked to report their own education in each survey year and their fathers' (or stepfathers') education in 1979. We relied on respondents' highest education across survey years. The final analysis for educational mobility included 5,494 cases with valid data on male respondents and 5,386 female respondents and their fathers.

The data are publicly available and can be downloaded from the NLS website: <http://www.nlsinfo.org/investigator/pages/login.jsp>.

### **National Survey of Families and Households**

The National Survey of Families and Households (NSFH) was a nationally representative household survey designed to provide information on family life. It was conducted in 3 waves: 1987–1988, 1992–1994, and 2001–2002. Households were randomly selected from 1,700 selection units that resulted from 17 enumeration districts within each of the 100 primary sampling units. For the first wave of data collection, a primary respondent at least 19 years of age was randomly selected from each household to participate in a personal interview. We used data from waves 1 to 3 and applied sample weight (*WEIGHT*) to account for the multistage sampling design. We kept primary respondents aged 25 to 64 years old in each wave.

Respondents were asked to report their own occupations in each wave and their fathers' occupations when the respondent was 16 years old. In order to generate occupational percentile ranks, we calculated individuals' birth years using age records in every round of the survey and converted their occupations reported at the age closest to 40 across all the waves to standardized 1950 occupation codes and then to microclass occupations. Father's birth year was asked only if the respondent coresided with his or her father. We imputed missing values in father's birth year using the same imputation method described in the GSS data section. The final analysis for occupational mobility relied on 3,373 male respondents and 3,935 female respondents with valid data and their reported fathers.

Respondents were asked to report their own education and their fathers' education when the respondent was 16 years old in the 1987–1988 survey. The final analysis for educational mobility included 3,048 male respondents and 4,074 female respondents.

The data are publicly available and can be downloaded from the NSFH website: <http://www.ssc.wisc.edu/nsfh/home2.htm>.

### **Occupational Changes in a Generation II**

Occupational Changes in a Generation II (OCG II) was collected in 1973 as a supplement to the March 1973 Current Population Survey. It was designed as a strict replication of OCG I, but also incorporated some new questions about social background and career development. The target population of OCG II was US males aged 20 to 64 in the civilian, noninstitutional population, except that household wives were also added to the sample. The resulting sample represented 95.4% of men and women aged 20 to 64. We applied the sample weight (*V582*) to account for the multistage sampling design and kept respondents aged 25 to 64.

Respondents were asked to report their own occupations at the time of the survey and their fathers' occupations when the respondent was 16 years old. We converted all occupations to the standardized 1950 occupation codes and then to microclass occupations. Because father's birth year was not asked in the OCG II data, we imputed the values using the same imputation method described in the GSS data section. The final analysis for occupational mobility relied on 20,350 cases with valid data on male respondents and 10,188 female respondents, and their fathers.

Individuals were asked to report their own education at the time of the survey and their fathers' education when the respondent was 16 years old. Unlike information on occupation, household wives were not asked about their education. The final analysis for educational mobility relied on 20,340 cases with valid data on male respondents and their fathers.

The data are publicly available and can be downloaded from the OCG website: <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/6162>.

### **Panel Study of Income Dynamics**

The Panel Study of Income Dynamics (PSID), carried out at the Survey Research Center at the University of Michigan, is a longitudinal panel household survey in the US. The survey collects extensive economic, social, and health items on individuals and their household members over the life course and across generations. Begun in 1968, the PSID follows over 18,000 household members from roughly 5,000 US families. The study interviewed individuals annually until 1997 and biennially thereafter. All original 1968 PSID respondents and their offspring are considered to carry the PSID "gene" and thus are permanent PSID respondents. Their demographic and socioeconomic information, such as age and occupation, is gathered in each wave of the PSID survey and can be linked across years. The PSID project also provides a "Family Identification Mapping System" (FIMS) tool designed to link family members across generations. The data consist of 2 distinct samples: SRC (Survey Research



Center) is a nationally representative household sample based on a stratified, multistage selection of the civilian noninstitutional population of the US; SEO (Survey of Economic Opportunity) is a national sample of low-income families with household heads under age 60 in 1968. We used data from year 1968 to 2017 in the SRC sample. The final analysis was restricted to male respondents aged 25 to 64.

Occupation is collected for household heads and their spouses in all PSID waves. We first harmonized all the occupation variables by converting them to the 1950 Census scheme and then chose individuals' occupations measured in the year closest to age 40 as the lifetime occupation. That is, if occupation at age 40 was unavailable, we then looked for occupation at ages 41, 39, 42, 38, and so forth. We converted all occupations to the standardized 1950 occupation codes and then to microclass occupations. For individuals whose fathers were not PSID respondents and thus were not directly observed, we relied on the question that asked household heads to report father's occupation in waves 1997–2017. If father's occupation was reported multiple times, we chose the mode of the data values. For missing data on father's birth year, we used the same imputation method described in the GSS data section. The final analysis for occupational mobility included 5,950 cases with valid data on father–son dyads and 2,321 father–daughter dyads from the years 1968–2017.

Level of education was asked for household heads in all PSID waves. For individuals whose fathers were not PSID respondents and thus were not directly observed, we relied on the question that asked household heads to report father's education. We chose the highest level of education for individuals and their fathers across waves. The final analysis for educational mobility included 8,765 cases with valid data on father–son dyads and 4,181 father–daughter dyads from years 1968–2017.

The data are publicly available and can be downloaded from the PSID website: <https://simba.isr.umich.edu/data/data.aspx>.

### **Survey of Income and Program Participation**

The Survey of Income and Program Participation (SIPP), conducted by the United States Census Bureau, is a longitudinal panel survey that is designed to provide income of individuals and households and their participation in income transfer programs. Other topics include education, occupation, family dynamics, health insurance, childcare, and food security. The SIPP survey design is a continuous series of national panels, with the sample size ranging from approximately 14,000 to 52,000 interviewed households. The original goal was to have all panels participate for a 32-month period (8 waves), with each panel randomly divided into one of 4 rotation groups. Each rotation group is interviewed in a separate month. Four rotation groups constitute one wave of interviewing. At each interview, respondents provide information covering the 4 months since the previous interview. The first interview began in October 1983 with the 1984 panel. Subsequent panels began interviews in February of each year. We used data from survey years 1986, 1987, and 1988 because the information about fathers' occupations and education was collected only in the wave 2 topical module during these years. We applied the sample weight (*FNLWGT\_5*) to account for the multistage sampling design and kept respondents aged 25 to 64.

Respondents were asked to report their own occupations during each survey year as well as their fathers' occupations when the respondent was 16 years old. We converted all occupations to the standardized 1950 occupation codes and then to microclass occupations. Because father's birth year was not asked in the SIPP surveys, we imputed the values using the same imputation method described in the GSS data section. The final analysis for occupational mobility relied on 14,547 cases with valid data on male respondents and 14,693 cases on female respondents and their fathers.

Respondents were asked to report their own education during each survey year as well as their fathers' education when the respondent was 16 years old. The final analysis for educational mobility relied on 13,644 cases with valid data on male respondents and 15,641 female respondents and their fathers.

The data are publicly available and can be downloaded from the SIPP website:

<http://www.census.gov/sipp/>.

## **S5 Occupation Measure in China**

Occupational categories were not measured consistently across censuses and surveys in our China data. We harmonized occupational variables from different data sources into the 2-digit occupational classification of the 2000 China Census. For data from the 2000 China Census and the 2005 China Mini-Census, occupations have already been coded using the 2000 Census classification. For data from the 1980 and 1990 China Censuses, we mapped occupations in the 1980 and 1990 Census classifications to the 2000 Census scheme. For data from LHSCCC1996 and CGSS2005–2018, occupations were coded using the International Standard Classification of Occupations (ISCO) 1968 and 1988, respectively. We similarly mapped the ISCO occupational classifications to the 2000 Census scheme. Containing 71 unique occupational categories, the 2000 Census scheme constitutes a basis for constructing our rank-based measure of occupational status.

## **S6 Occupation Measure in the US**

We used the occupation variable OCC1950 in the IPUMS US Census/ACS data. All occupations were coded into the 1950 Census Bureau occupational classification. The documentation of the 1950 classification can be found in “Integrated Occupation and Industry Codes and Occupational Standing Variables in the IPUMS” (<https://usa.ipums.org/usa/chapter4/chapter4.shtml>) and the “Alphabetic Index of Occupations and Industries: 1950” (<https://usa.ipums.org/usa/resources/volii/Occupations1950.pdf>). The original 1950 occupational classification consists of 283 occupational categories, but occupations with codes above 970 are excluded from our analysis (979 Not yet classified; 980 Keeps house/housekeeping at home/housewife; 981 Imputed keeping house; 982 Helping at home/helps parents/housework; 983 At school/student; 984 Retired; 985 Unemployed/without occupation; 986 Invalid/disabled w/ no occupation reported; 987 Inmate; 990 New worker;

991 Gentleman/lady/at leisure; 995 Other nonoccupational; 997 Occupation missing/unknown; and 999 N/A).

The occupational measures used in the final analysis were generated by means of the following steps. We first mapped the OCC1950 to OCC1960 using a crosswalk file and then mapped OCC1960 to the microclass occupational scheme for the US used in a study by Jonsson *et al.* (2). For microclass occupations that required special coding, such as additional information about self-employment status (class of worker), we dropped these categories because of the limited information collected in the historical census data. Specifically, we excluded individuals coded as 9999 Occupation not reported. We then dropped several occupations, 1301 Systems analysts and programmers, 2001 Proprietors, 4202 Chemical processors, and 1314 Nursery school teachers and aides, because we were unable to identify workers in these occupations based on the crude information of OCC1950. We coded individuals who potentially worked in these occupations into other related occupations. We then collapsed microclass occupations that were not consistently observed with other closely related occupations. Specifically, we grouped 1105 Statistical and social scientists into 1104 Natural scientists; 1302 Aircraft pilots and navigators into 1308 Professional, technical, and related workers, n.e.c.; 1303 Personnel and labor relations workers into 1308 Professional, technical, and related workers, n.e.c.; 1309 Social and welfare workers into 1308 Professional, technical, and related workers, n.e.c.; 4106 Electricians into 4103 Electronics service and repair workers; 4108 Vehicle mechanics into 4111 Other mechanics; 4120 Heavy machine operators into 4201 Truck drivers; 4204 Longshoremen and freight handlers into 4209 Operatives and kindred workers, n.e.c.; and 4205 Food processors into 4209 Operatives and kindred workers, n.e.c. Finally, we created a crosswalk file that mapped OCC1950 to 70 microclass categories detailed enough to capture heterogeneity among occupations and generate continuous occupational percentiles while also containing enough cases within each census year for a historical comparison. The same method was used in the study by Song *et al.* (1).

## **S7 Occupational Percentile Ranks**

We use a method that converts discrete measures of occupations into percentile ranks (1). Compared with education or income, occupational status is more difficult to measure because occupational categories lack an intrinsic scale. To generate ordered occupations, Otis Dudley Duncan developed a “socioeconomic index of occupations”—now known as the “Duncan SEI score”—that has been widely used in the literature and is available in the IPUMS US Census data. Duncan used prestige ratings (from the 1947 National Opinion Research Center study) as a way of obtaining weights for occupation education and occupation income and created a predicted SEI score for each occupation in the 1950 Census (3). We devised a new occupational ranking procedure that is closely related to Duncan’s method but can be used to analyze historical and cross-national data. Below, we expound the steps used to construct the occupational ranks for the US data. Occupational ranks for the Chinese data are constructed using the same procedure with the 71 occupational categories of the 2000 China Census.

First, we create a population-based occupational database aggregated by birth cohort, occupation, and education. Specifically, we draw on data from the IPUMS population censuses and the ACS from 1850 to 2015 and restrict the data to workers aged 25 to 64. All the occupational variables in the IPUMS data have been coded into the 1950 Census occupation categories. The data are grouped by birth cohort because occupational percentile ranks are assumed to vary over cohorts. For example, the occupational status of telephone operators would be higher in the 1900 birth cohort than in the 1970 cohort, as the telephone was still considered a new technology in the early 20th century.

Second, we harmonize occupational measures across years using a revised version of Grusky *et al.*'s microclass occupational scheme (2, 4, 5), with the 269 occupational categories in the 1950 Census mapped to the 70 categories of microclass occupations (1). This mapping step results in a dataset that contains 1,400 observations, each of which refers to a microclass occupation for a certain birth cohort that falls within the range of 1790 to 1980. Other variables in the dataset include the number of workers within each occupation and the number of persons with varying levels of education (0, 1–8, 9–10, 11, 12, 13–15, 16 and more years of schooling). The detailed education variable was not available until the 1940 Census. For years prior to 1940, we generate occupations' literacy scores from a dichotomous variable (0 = illiterate; 1 = literate, can both read and write).

Third, we create occupational status based on the educational distribution within each occupation. For occupation  $i$ , its status score is

$$S_{it} = \sum_x p(x|i, t) \cdot Q_t(r^x)$$

where  $p(x|i, t)$  is the proportion of educational level  $x$  in occupation  $i$  and birth cohort  $t$ ;  $Q_t(r^x)$  is the percentile rank of educational level  $x$  in birth cohort  $t$ . For example, assume we have 4 educational groups that are ranked from 1 (low) to high (4) and vary in size from 40, 30, 20, to 10 in a general population that contains 100 individuals in total. The percentile rank of group 4 is 95; namely, the midpoint of the 90th percentile and the 100th percentile because of tied values within this educational group. Likewise, the percentile ranks of groups 1, 2, and 3 would be 20, 55, and 80, respectively. Assume that for a specific occupation  $i$ , the proportions of educational groups from 1 to 4 are 0.1, 0.35, 0.3, and 0.25, respectively. Thus, this occupation's status score is 69 ( $= 0.1 \cdot 20 + 0.35 \cdot 55 + 0.3 \cdot 80 + 0.25 \cdot 95$ ).

Overall, an occupation with more college-educated workers would have a higher status than an occupation with fewer college-educated workers, all other things being equal. However, because of the expansion of higher education, the status of the college-educated group per se has also evolved. An occupation with 20 percent college-educated workers in 1940 would have a higher status than an occupation with the same proportion of college-educated workers in 2000, all other things being equal. In other words, the relative percentile ranks of different educational groups have also changed over time because of the evolution of the educational distribution.

Fourth, we rank microclass occupations from 1 to 70 by their status scores,  $S_{it}$ , within each birth cohort. Because this rank is similar to Treiman’s international socioeconomic index of occupational status (6), we refer to it as “Treiman’s rank” in the following sections.

Fifth, we convert Treiman’s rank into percentiles from 0 (lowest) to 100 (highest), after accounting for variations in sizes among occupations and over birth cohorts. The percentile ranks are less stable over time, compared to Treiman’s ranks, because the former capture changes in the size of occupations. If a privileged occupation doubles in size without changes in the educational composition and thus increases its share in the overall population, then the status of this occupation would decrease because more workers would be tied in their statuses.

We use the moving average method to smooth out fluctuations caused by small  $N$ s. The adjusted percentile for the birth cohort  $t$  is

$$\text{pcrank\_adj}[t] = 0.25 * \text{pcrank}[t-1] + 0.5 * \text{pcrank}[t] + 0.25 * \text{pcrank}[t+1]$$

For birth cohorts at the 2 ends, when  $\text{pcrank}[t-1]$  or  $\text{pcrank}[t+1]$  is missing, we use

$$\text{pcrank\_adj}[t] = 0.25 * \text{pcrank}[t-1] + 0.75 * \text{pcrank}[t]$$

and

$$\text{pcrank\_adj}[t] = 0.75 * \text{pcrank}[t] + 0.25 * \text{pcrank}[t+1]$$

Sixth, we link occupational percentiles generated from the population data to the individual-level mobility table, based on father’s and child’s birth cohorts and occupations.

## S8 Educational Percentile Ranks

We use a method similar to that for constructing occupation percentile ranks (see *SI Appendix*, Subsection S7), which in this case converts discrete measures of educational levels into percentile ranks. The method proceeds in 4 steps.

First, for each country we create a population-based educational database aggregated by birth cohort, gender, and education. Specifically, for the US, we draw on data from the IPUMS population censuses and the ACS from 1940 to 2015 and restrict the data to men and women aged 25 and above. For China, we draw on data from IPUMS microdata samples of population censuses (1982, 1990 and 2000) and aggregated statistics of the 2010 Census. The data are grouped by birth cohort and gender.

Second, we calculate the percentile rank for each educational level according to the educational distribution specific to each gender and birth cohort. For example, assume we have 4 educational groups that are ranked from 1 (low) to high (4) and vary in size from 40, 30, 20, to 10 in a birth cohort that contains 100 men in total. The percentile rank of group 4 is 95; namely, the midpoint of the 90th percentile and the 100th percentile because of tied values within this educational group. Likewise, the percentile ranks of groups 1, 2, and 3 would be 20, 55, and 80, respectively. Due to the expansion of education and women’s greater

improvement in educational attainment relative to men's over time in both countries, educational distribution has greatly changed over cohorts and by gender. Hence, the relative percentile ranks of different educational groups have changed over time and differed between men and women.

Third, we determine percentile ranks to be used in the analysis for cohorts that are covered by multiple census waves and therefore have varying relative ranks across waves. For Chinese cohorts, we have just a few census waves. To allow most individuals to complete their education and avoid survival selection by education, we pick the percentile ranks calculated from the census wave when the relevant cohort were closest to the age 30–40. For US cohorts, since multiple census waves are available for each cohort, we take the average ranks across censuses if multiple data points are available for a cohort.

Fourth, we link educational percentile ranks generated from the population data to individuals and their fathers in the analytical samples, based on their gender, birth cohort, and country.

TABLE S1. INTERGENERATIONAL RANK–RANK CORRELATION IN OCCUPATION BY OFFSPRING’S BIRTH COHORT IN THE US AND CHINA

Offspring’s Birth Cohort	Father–Son Dyads				Father–Daughter Dyads			
	Full Population		Non-Agricultural Population		Full Population		Non-Agricultural Population	
	China	US	China	US	China	US	China	US
1946–1955	0.462	0.302	-0.033	0.291	0.537	0.267	-0.033	0.260
	[0.437, 0.486]	[0.287, 0.317]	[-0.084, 0.018]	[0.275, 0.307]	[0.512, 0.562]	[0.252, 0.283]	[-0.087, 0.021]	[0.244, 0.276]
1956–1965	0.421	0.329	0.038	0.322	0.507	0.265	0.032	0.264
	[0.399, 0.443]	[0.313, 0.345]	[0.000, 0.076]	[0.305, 0.338]	[0.487, 0.527]	[0.249, 0.281]	[-0.004, 0.069]	[0.247, 0.281]
1966–1975	0.336	0.304	0.034	0.295	0.433	0.261	0.116	0.259
	[0.314, 0.359]	[0.271, 0.336]	[-0.006, 0.074]	[0.261, 0.329]	[0.413, 0.453]	[0.226, 0.296]	[0.078, 0.155]	[0.222, 0.295]
1976–1985	0.374	0.337	0.181	0.330	0.427	0.267	0.088	0.258
	[0.346, 0.403]	[0.298, 0.376]	[0.136, 0.225]	[0.290, 0.370]	[0.399, 0.454]	[0.221, 0.312]	[0.045, 0.132]	[0.211, 0.304]

*Data sources:* GSS 1972–2018; NLS–YM 1966–1981; NLS–YW 1968–1993; NLSY 79, 1979–2012; NSFH, 1987, 1993, 2002; OCG II; PSID (SRC sample) 1968–2017; SIPP 1986, 1987, 1988; LHSCCC 1996; CGSS 2005, 2006, 2008, 2010, 2012, 2013, 2015, 2017, 2018.

*Notes:* Numbers in parentheses represent 95% confidence intervals.

TABLE S2. AVERAGE OCCUPATIONAL PERCENTILE RANK FOR MEN AND WOMEN BY BIRTH COHORT, FATHER'S OCCUPATION, AND *HUKOU* ORIGIN IN CHINA

Cohort	Father's Occupation	All		Rural <i>Hukou</i> Origin		Urban <i>Hukou</i> Origin	
		Men	Women	Men	Women	Men	Women
1946–55	Farmer/Farm Laborer	57.23	52.69	55.90	50.93	80.20	79.91
	Manual Worker	80.73	80.37	71.06	65.66	84.60	85.03
	Nonmanual Worker	79.55	79.25	68.68	63.68	87.61	89.01
1956–65	Farmer/Farm Laborer	55.91	50.75	54.33	49.18	77.43	77.36
	Manual Worker	78.18	77.76	71.30	64.78	80.66	81.89
	Nonmanual Worker	78.35	78.29	69.84	65.95	84.62	86.51
1966–75	Farmer/Farm Laborer	58.00	53.18	56.59	51.79	78.37	76.44
	Manual Worker	75.62	76.28	71.13	68.00	79.07	83.09
	Nonmanual Worker	76.74	79.40	70.16	71.31	83.03	88.20
1976–85	Farmer/Farm Laborer	59.29	60.15	58.03	59.13	74.52	75.00
	Manual Worker	73.79	78.77	71.52	74.89	75.77	82.73
	Nonmanual Worker	77.36	80.31	74.45	74.48	79.46	85.52

*Data sources:* LHSCCC1996; CGSS 2005, 2006, 2008, 2010, 2012, 2013, 2015, 2017, 2018.



TABLE S3. SELECTED MAJOR OCCUPATION BY GENDER AND *HUKOU* ORIGIN IN CHINA

Occupation Destination	Cohort	Rural <i>Hukou</i> Origin		Urban <i>Hukou</i> Origin	
		Men	Women	Men	Women
Farmers/Farm Laborers	1946–55	0.655	0.772	0.038	0.031
	1956–65	0.528	0.621	0.030	0.028
	1966–75	0.402	0.471	0.022	0.024
	1976–85	0.275	0.321	0.014	0.020
Managerial/Professional Workers/Large Proprietors	1946–55	0.090	0.033	0.251	0.237
	1956–65	0.120	0.076	0.299	0.301
	1966–75	0.151	0.120	0.347	0.365
	1976–85	0.182	0.163	0.396	0.429

*Data sources:* LHSCCC1996; CGSS 2005, 2006, 2008, 2010, 2012, 2013, 2015, 2017, 2018.

TABLE S4. INTERGENERATIONAL RANK–RANK CORRELATION IN EDUCATION BY OFFSPRING’S BIRTH COHORT IN THE US AND CHINA

Offspring’s Birth Cohort	Father–Son Dyads		Father–Daughter Dyads	
	China	US	China	US
1946–1955	0.243 [0.210, 0.276]	0.409 [0.394, 0.424]	0.361 [0.328, 0.394]	0.415 [0.401, 0.430]
1956–1965	0.322 [0.297, 0.346]	0.434 [0.418, 0.450]	0.355 [0.329, 0.381]	0.416 [0.400, 0.432]
1966–1975	0.383 [0.360, 0.405]	0.456 [0.423, 0.488]	0.430 [0.409, 0.452]	0.428 [0.394, 0.461]
1976–1985	0.397 [0.371, 0.424]	0.423 [0.384, 0.461]	0.447 [0.422, 0.473]	0.403 [0.364, 0.443]

*Data sources:* GSS 1972–2018; NLS–YM 1966–1981; NLS–YW 1968–1993; NLSY79 1979–2012; NSFH 1987; OCG II (only males); PSID (SRC sample) 1968–2017; SIPP 1986, 1987, 1988; LHSCCC 1996; CGSS 2005, 2006, 2008, 2010, 2011, 2012, 2013, 2015, 2017, 2018.

*Notes:* Numbers in parentheses represent 95% confidence intervals.

TABLE S5. AVERAGE YEARS OF SCHOOLING FOR MEN AND WOMEN BY BIRTH COHORT, FATHER'S EDUCATION, AND *HUKOU* ORIGIN IN CHINA

Cohort	Father's Education	All		Rural <i>Hukou</i> Origin		Urban <i>Hukou</i> Origin	
		Men	Women	Men	Women	Men	Women
1946–55	Primary School or Lower	7.97	6.04	7.33	4.68	9.78	9.20
	Junior/Senior Middle School	10.14	9.57	8.82	6.45	10.87	11.14
	Junior College or Higher	11.65	11.47	–	–	12.35	11.92
1956–65	Primary School or Lower	9.39	7.77	8.75	6.57	10.90	10.59
	Junior/Senior Middle School	11.47	10.42	10.45	8.24	12.16	11.83
	Junior College or Higher	12.87	12.58	11.45	10.63	13.38	13.27
1966–75	Primary School or Lower	9.11	7.82	8.71	7.18	10.98	10.76
	Junior/Senior Middle School	11.50	10.81	10.58	9.58	12.67	12.43
	Junior College or Higher	14.15	13.91	13.03	12.01	14.50	14.53
1976–85	Primary School or Lower	10.20	9.21	9.80	8.79	12.04	11.61
	Junior/Senior Middle School	12.79	12.39	11.81	11.09	13.94	14.22
	Junior College or Higher	15.53	15.27	14.56	13.97	15.87	15.60

*Data sources:* LHSCCC1996; CGSS 2005, 2006, 2008, 2010, 2011, 2012, 2013, 2015, 2017, 2018. Due to insufficient data, the average years of schooling are not computed for men and women in the 1946–55 birth cohort who had a rural *hukou* origin and a college-educated father.

## **Bibliography**

DUNCAN, O. D., “Chapter 6: a socioeconomic index for all occupations” and “Chapter 7: properties and characteristics of the socioeconomic index” in *Occupations and social status*, A. J. Reiss, Jr., Ed. (New York: Free Press of Glencoe, 1961), 109–138, 139–161.

JONSSON, J. O., *et al.*, Microclass mobility: Social reproduction in four countries. *Am. J. Sociol.* 114, 977–1036 (2009).

SONG, X., *et al.*, Long-term decline in intergenerational social mobility in the United States, 1850–2015. *Proc. Natl. Acad. Sci. U.S.A.* 117, 251–258 (2020).

TREIMAN, D. J., *Occupational prestige in comparative perspective* (New York: Academic Press, 1977).

WEEDEN, K. A., GRUSKY, D. B., The case for a new class map. *Am. J. Sociol.* 111, 141–212 (2005).

WEEDEN, K. A., GRUSKY, D. B., The three worlds of inequality. *Am. J. Sociol.* 117, 1723–1785 (2012).



**Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex) distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR). Il est financé dans le cadre des investissements d'avenir (ANR-11-LABX-0091, ANR-11-IDEX-0005-02) et de l'IdEx Université de Paris (ANR-18-IDEX-0001).**

**[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)**

## **A propos de la publication**

### **Procédure de soumission :**

Rédigé par un ou plusieurs chercheurs sur un projet en cours, le *Working paper* vise à susciter la discussion scientifique et à faire progresser la connaissance sur le sujet étudié. Il est destiné à être publié dans des revues à comité de lecture (peer review) et à ce titre répond aux exigences académiques. Les textes proposés peuvent être en français ou en anglais. En début de texte doivent figurer : les auteurs et leur affiliation institutionnelle, un résumé et des mots clefs.

Le manuscrit sera adressé à : [liepp@sciencespo.fr](mailto:liepp@sciencespo.fr)

Les opinions exprimées dans les articles ou reproduites dans les analyses n'engagent que leurs auteurs.

### **Directrice de publication :**

Anne Revillard

### **Comité de rédaction :**

Sofia Cerda Aparicio, Andreana Khristova

Sciences Po - LIEPP  
27 rue Saint Guillaume  
75007 Paris - France  
+33(0)1.45.49.83.61  
[liepp@sciencespo.fr](mailto:liepp@sciencespo.fr)

