NBER WORKING PAPER SERIES

WORKPLACE SEGREGATION BETWEEN COLLEGE AND NON-COLLEGE WORKERS

Francis M. Dillon
Edward L. Glaeser
William R. Kerr

Working Paper 33653
http://www.nber.org/papers/w33653

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2025

We thank seminar participants and panelists at the American Economic Association meetings for helpful comments. Lydia Qin provided excellent research assistance. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2766. (CBDRB-FY24-P2766-R11646, R11892, R11898, FY25-P2766-R12120). This research uses data from the Census Bureau's Longitudinal Employer-Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w33653

Workplace Segregation between College and Non-college Workers
Francis M. Dillon, Edward L. Glaeser, and William R. Kerr
NBER Working Paper No. 33653
April 2025
JEL No. I24, J24, J31, J62, L23, O15

ABSTRACT

We measure the level and growth of education segregation in American workplaces from 2000 to 2020. American workplaces show an educational segregation, measured by the degree to which the establishment has mostly workers of similar education levels, that is comparable to racial residential segregation in a typical metro area. Workplace isolation was particularly high for young and male workers without college degrees. The isolation of non-college workers is increasing over time.

Francis M. Dillon
Harvard University
fdillon@hbs.edu

Edward L. Glaeser
Department of Economics
315A Littauer Center
Harvard University
Cambridge, MA 02138
and NBER
eglaeser@harvard.edu

William R. Kerr
Harvard Business School
Rock Center 212
Soldiers Field
Boston, MA 02163
and NBER
wkerr@hbs.edu

Introduction

How segregated are American workplaces in terms of education and skills? For sixty years, social scientists have researched residential segregation (Taeuber and Taeuber, 1965; Murray, 2012; Athey et al., 2021), but the typical employed American spends about eight hours at work on a normal workday. This time at work is far more than the less than one hour per day that adults usually spend "socializing and communicating" and in "organization, civic and religious activities."[1]

Moreover, events at work seem far more likely to shape long-term skill acquisition by workers and advance their economic outcomes than neighborhood events. Indeed, the level of human capital in the workplace may beget more human capital and productivity in the future (e.g., Becker, 1964; Lucas, 1998; Duranton and Puga, 2000; Moretti, 2004; Mas and Moretti, 2009).

We measure the level of educational segregation in US workplaces experienced by non-college workers using the Longitudinal Employer-Household Dynamics (LEHD) data. During the 2000-2020 period, educational dissimilarity of the workplace was comparable in magnitude to neighborhood racial segregation for housing and rose somewhat. Workplace isolation was particularly high for young and male workers without college degrees.

A companion paper, Dillon, Glaeser and Kerr (2025), finds that workplace isolation negatively impacts the future careers of non-college workers who are less exposed to college-educated workers. Barzu et al. (2024) also find a strong correlation between workplace segregation and economic success in Brazil. The growing educational segregation of our workplaces may therefore limit the ability of younger, less-skilled workers to learn on the job from their more-educated peers, diminishing acquisition of skills that promote higher-paid future employment (e.g., Blair et al., 2020).

I. National Trends

The LEHD is built from unemployment insurance records and covers around 95% of all private sector workers (US Census Bureau 2000-2020). Our project has access to 22 states with

[1] Data from the 2019 American Time Use Survey (https://www.bls.gov/news.release/archives/atus_06252020.pdf).

full records for 1998-2020. Our states include California and Pennsylvania and comprise slightly less than one-half of the American population. We include all workers aged 20-64 in each year where we have non-imputed education data.[2] Workers must be present at least two quarters where they earn $2000 or more (in inflation-adjusted 2000 dollars) and have one job that accounts for the majority of their earnings. We organize the data around work-block locations of firms, which we term "establishments" in this paper. Dillon, Glaeser and Kerr (2025) provides an extended data description.

Table 1 considers the level of education and the workplace segregation by education over time. The first column shows the well-known fact that American workers are becoming more educated. Sixty-eight percent of workers in our sample lacked a college degree in 2000. By 2020, that share had fallen to sixty-two percent.

Table 1: Trends in workplace education levels and educational segregation

	Share of non-college workers	Isolation index, non-college	Adjusted isolation index, non-college	Isolation index, college	Adjusted isolation index, college	Dissimilarity index
	(1)	(2)	(3)	(4)	(5)	(6)
2000	0.679	0.787	0.108	0.549	0.228	0.519
2010	0.643	0.768	0.126	0.583	0.226	0.525
2020	0.620	0.767	0.147	0.620	0.240	0.550
Δ 2000-2020	-0.059	-0.020	0.039	0.071	0.012	0.031

Notes: Table documents trends in workplace education levels and educational segregation using our LEHD dataset with 22 states present from 2000 to 2020. Included individuals are aged 20-64 and working in a private-sector establishment with at least nine other employees. The sample is limited to individuals with non-imputed education values. Workers must be present at least two quarters where they earn $2000 or more (in inflation-adjusted 2000 dollars) and hold one job that accounts for the majority of their earnings. Column 1 presents the share of workers in LEHD without a college degree among this sample. Column 2 presents the isolation index for non-college workers. Column 3 provides the isolation index that adjusts for workforce shares of education levels. Columns 4 and 5 repeat these metrics for college-educated workers. Column 6 provides the dissimilarity index. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

One measure of educational segregation is the isolation index (Taeuber and Taeuber, 1964):

$$(1)\ Isolation\ Index = \sum_{Establishments} \frac{Group_{Est}}{Total_{Est}} \frac{Group_{Est}}{Group_{Total}}$$

where $Group_{Est}$ is the number of non-college workers in the establishment, $Total_{Est}$ is the total number of workers in the establishment, and $Group_{Total}$ is the total size of the non-college workforce across all establishments. This index measures the share of non-college workers in the

[2] The appendix provides an extended discussion on the implications of focusing on non-imputed data and documents alternative series with empirical adjustments for potential biases. The series for non-college workers are very similar with adjustments, but the college workers series are sensitive.

establishment of the average non-college worker. Column 2 shows that this index has declined from 0.787 in 2000 to 0.767 in 2020, suggesting that non-college workers are exposed to slightly fewer non-college workers than in the past. This is unsurprising given the overall decline in the share of non-college workers in Column 1.

The third column reports an adjusted isolation index, which accounts for this aggregate trend by subtracting the overall population share of the group nationally:

(1') $Adjusted\ Isolation\ Index = Isolation\ Index - \frac{Group_{Total}}{Total_{Total}}$,

where $Total_{Total}$ is the total population of workers. The adjusted isolation index subtracts the total share of non-college workers in the LEHD sample from the basic isolation index and shows an increase over time from 0.108 to 0.147. Non-college workers are being exposed to more college-educated workers over time, but the increase in exposure is less than what one might expect given the overall increase nationally in education levels.

Columns 4 and 5 provide similar workplace isolation indices for college-educated workers. The adjusted isolation index for college workers starts at 0.228 in 2000, more than twice the initial level of non-college workers. Adjusted isolation for college-educated workers also rises to a value of 0.240 by 2020, but the increase is less than a third of the increase for non-college workers.[3]

Column 6 presents an aggregated view measured by a dissimilarity index, defined as

(2) $Dissimilarity\ Index = \frac{1}{2}\sum_{Establishments}\left|\frac{Group_{Est}}{Group_{Total}} - \frac{Non-Group_{Est}}{Non-Group_{Total}}\right|$,

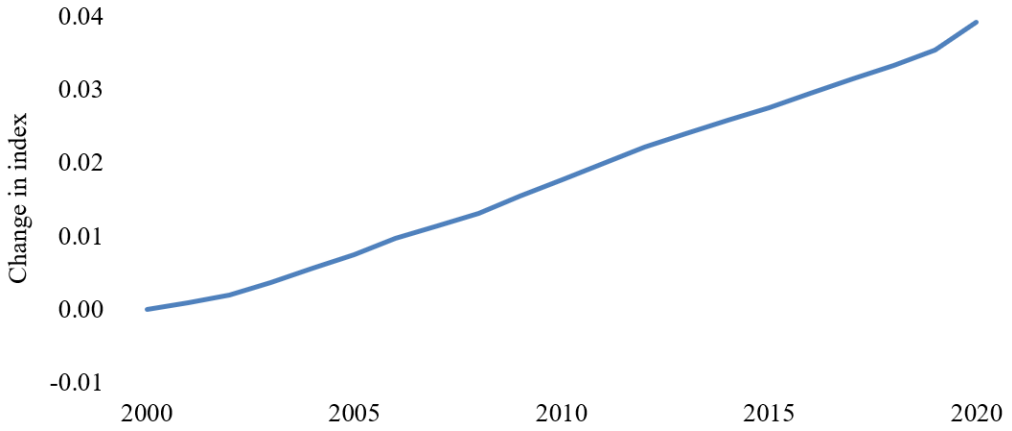
where $Non - Group_{Est}$ is the number of college-educated workers in the establishment and $Non - Group_{Total}$ is the total number of college-educated workers in the workforce. This measure represents the share of non-college workers (or, equivalently, college-educated workers) who would need to move across establishments to create an even distribution of education across

[3] These two series need not mirror each other. For example, consider an economy with three firms, two perfectly segregated by education and the other mixed. If the non-college workforce grows by adding to the education-segregated firm, and the college-educated workforce grows by adding to the education-mixed firm, then the isolation of the non-college workers can increase while the isolation of college workers falls.

workplaces. The index rose from 0.519 to 0.550 between 2000 and 2020. These figures are comparable to the 0.53 racial dissimilarity index for the median large metropolitan area in 2020.[4]

Figure 1 shows the change in the adjusted isolation index of non-college workers since 2000. The growth in the isolation of non-college workers has been consistent during the period.

Figure 1: Change in adjusted isolation index for non-college workers, 2000-2020



Notes: Figure plots change in workplace isolation for non-college workers from the starting levels in 2000. Table 1 describes the underlying sample. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

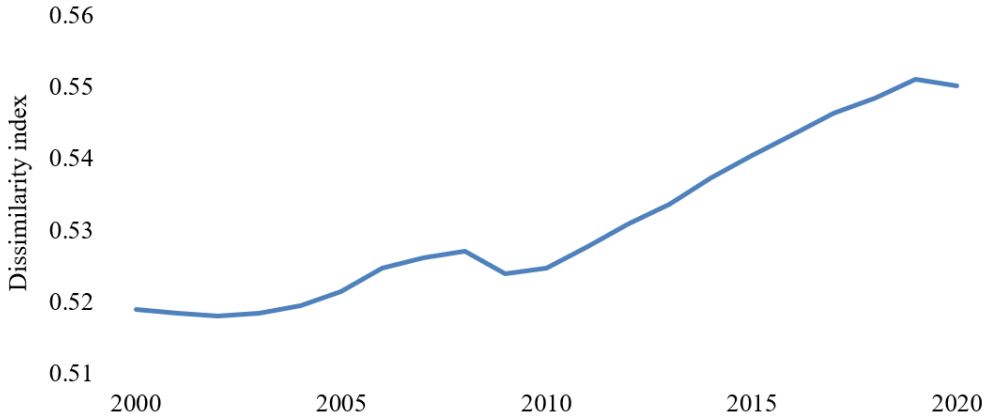
The growth in the dissimilarity index, which is the same for both education groups, shows a less steady pattern in Figure 2. The index rose modestly from 2000 to 2010, rising before the global financial crisis and then dipping after. But between 2010 and 2019, the index's rise was steady. There was a down tick in dissimilarity in 2020, but it is too early to know whether this is a shift in the long-term trajectory or will prove to be temporary.

An increase in workplace segregation can happen because of shifts in employment towards more-segregated workplaces, an increase in establishment-level segregation, or both. The appendix shows that the average level of educational segregation within firms is increasing over time. While small firms have the highest overall level of educational segregation, the increase since 2000 is evident among small, medium and large firms. In some cases, segregation has increased

[4] City Observatory: America's least (and most) segregated metro areas: 2020. (https://cityobservatory.org/most_segregated2020/).

because firms have outsourced specialized operations typically performed by less-educated workers, such as janitorial services, food services, and security.

Figure 2: Dissimilarity index, 2000-2020



Notes: Figure plots workplace dissimilarity index from 2000 to 2020. The index measures the share of the non-college workforce (or, equivalently, the college-educated workforce) that would need to move establishments in order to produce a uniform education share across workplaces. Table 1 describes the underlying sample. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

II. Heterogeneity for Non-College Workers

Table 2 looks at the changes in workplace segregation for non-college workers by demographic and regional subgroups. Men had higher adjusted isolation and dissimilarity indices than women, and they have experienced greater growth in these indices. The patterns by race are more mixed. White non-college workers were in more skill-segregated workplaces than minorities in 2000; the measures disagree about which group's skill isolation increased more since then. Across age groups, young non-college workers experience the most workplace isolation, and this age gap has been growing.

The bottom four rows of Table 2 show that workplace segregation is highest and growing most quickly in the West. Segregation is lowest in the Northeast. These patterns stand in contrast with racial residential segregation levels, which are highest in old large metropolitan areas and lowest in western cities (Cutler, Glaeser and Vigdor, 1999). One potential explanation for this divergence is that racial residential segregation has been declining over time, and consequently it is lowest in our newest communities which are less trapped in the past. Conversely, workplace

segregation by education level is rising and may be highest in cities and industries least likely to reflect legacies of the past (e.g., Glaeser, Kerr and Kerr, 2015).

Table 2: Trends in workplace isolation for non-college workers by worker type

	Adjusted isolation index, 2000 (1)	Dissimilarity index, 2000 (2)	Δ Adjusted isolation index, 2000-2020 (3)	Δ Dissimilarity index, 2000-2020 (4)
Total	0.108	0.519	0.039	0.031
Men	0.141	0.523	0.046	0.042
Women	0.069	0.515	0.034	0.020
White	0.111	0.522	0.041	0.026
Minority	0.091	0.504	0.037	0.052
Young (18-29)	0.125	0.518	0.062	0.044
Middle Aged (30-49)	0.105	0.519	0.026	0.036
Older (50-64)	0.097	0.519	0.051	0.016
Northeast	0.091	0.503	0.029	0.022
Midwest	0.108	0.515	0.038	0.024
South	0.110	0.529	0.034	0.020
West	0.116	0.523	0.047	0.044

Notes: See Table 1. Table documents levels and changes in workplace isolation by worker type among non-college workers. For indices by worker demographic group, establishments are weighted by the number of people of indicated type. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

Table 3 shows the indices across classes of firms. In 2000, goods-producing firms were somewhat more integrated, by either measure, than services-producing firms. Moreover, the adjusted isolation index increased by less in goods-producing firms, although the dissimilarity index increased similarly in goods- and service-producing firms. Isolation and dissimilarity were both lower in multi-unit than in single-unit firms, suggesting complex organizations have more mixing of education groups. However, adjusted isolation and dissimilarity rose much more in multi-unit firms, which may reflect the outsourcing of specialized tasks or rise in superstar firms that hire particularly skilled workers (e.g., Kremer and Maskin, 1996; Song et al., 2019; Autor et al., 2020).

Small establishments were considerably more segregated by education than larger establishments in 2000. Measured segregation is typically larger in smaller units in part due to mechanical reasons. If a unit is a single person, then segregation is always perfect, and if a unit is

the entire population, then dissimilarity or adjusted isolation is always zero. Both of our measures of segregation also rose within small establishments. Similarly, both measures of educational segregation were higher in new firms in 2000 and the increase in segregation was bigger for new firms.

Table 3: Trends in workplace isolation for non-college workers by firm type

	Adjusted isolation index, 2000 (1)	Dissimilarity index, 2000 (2)	Δ Adjusted isolation index, 2000-2020 (3)	Δ Dissimilarity index, 2000-2020 (4)
Total	0.108	0.519	0.039	0.031
Goods-producing firm	0.048	0.482	0.025	0.036
Services-producing firm	0.121	0.503	0.038	0.035
Single-unit firm	0.141	0.602	0.042	0.018
Multi-unit firm	0.082	0.451	0.043	0.054
Small establishment	0.126	0.626	0.052	0.054
Medium establishment	0.082	0.461	0.024	0.022
Large establishment	0.082	0.366	0.027	0.015
New firm	0.157	0.630	0.056	0.073
Existing firm	0.102	0.506	0.041	0.035

Notes: See Table 1. Table documents levels and changes in workplace isolation by firm type among non-college workers. Small establishments have between 10 and 99 workers, medium establishments have between 100 and 999 workers, and large establishments have at least 1,000 workers. New firms are those founded between 1996-2000, and existing firms must be founded in 1995 or earlier. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

Table 4 shows the five industries contributing the most to the national increase in isolation. Workplace segregation rose due to the shift of non-college workers into extremely isolated industries. Four of the five industries that contributed most to rising isolation are not themselves becoming more isolated, but they account for a growing share of the non-college workforce (e.g., warehousing and storage). Restaurants contribute through both growing sector size and because their establishments are becoming more segregated. These industries explain 43 percent of the increase in adjusted isolation between 2000 and 2020.

Table 4: Major industries contributing to rise in adjusted isolation index

	Share of non-college workers, 2000 (1)	Δ Share of non-college workers, 2000-2020 (2)	Isolation index, 2000 (3)	Δ Isolation index, 2000-2020 (4)	Percent contribution to adjusted isolation index, 2000-2020 (5)
Restaurants and other eating places	3.68	0.43	0.872	0.055	13.9
Warehousing and storage	0.68	1.58	0.871	-0.031	9.3
General merchandise stores	0.98	1.01	0.887	0.000	8.3
Services to buildings and dwellings	1.20	0.53	0.917	-0.005	5.5
Grocery stores	2.02	0.39	0.887	-0.005	5.4

Notes: See Table 1. Table documents five major industries contributing to the rise in adjusted isolation index for non-college workers. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898.

III. Future Directions

The segregation of US workplaces between college and non-college worker is sizable and growing. Workplace isolation is particularly high for young and male non-college workers. In Dillon, Glaeser and Kerr (2025), we build off these trends to demonstrate how proximity to college workers in the workplace boosts career prospects for non-college workers. Our analysis looks at how proximity to college-educated workers in 2000 links to higher earnings and future growth for non-college workers over the next 20 years, and we consider several types of events (e.g., mass worker moves) that exogenously shock the workplace education levels surrounding non-college workers to confirm the role of the workplace versus unobserved traits of workers.

We also anticipate extending this work. By observing workers' changes in occupations over time, we can measure the types of skills being accumulated, such as abstract reasoning and social skills (e.g., Deming, 2017). We plan to integrate our work on educational segregation with race and gender dynamics in the workplace. The accumulation of skills for a non-college worker may be diminished if the college-educated workers in their workplace have different demographics. Finally, there is much to learn about how firms decide about establishment composition and outsourcing and the impact of these decisions on their workers.

References

Athey, S. Ferguson, B., Gentzkow, M., & Schmidt, T. (2021). Estimating experienced racial segregation US cities using large-scale GPS data. Proceeding of National Academy of Sciences 118 (46) e2026160118.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. The Quarterly Journal of Economics 135 (2): 645-709.

Barzu, R., Glaeser, E., Hidalgo, C., & Viarengo, M. (2024). Cities as engines of opportunity: Evidence from Brazil. NBER Working Paper.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago Press.

Blair, P.Q., Castagnino, T.G., Groshen, E.L., Debroy, P., Auguste, B., Ahmed, S., Garcia Diaz, F., & Bonavida, C. (2020). Searching for STARs: Work experience as a job market signal for workers without bachelor's degrees. NBER Working Paper No. 26844.

City Observatory. (2020). America's least (and most) segregated metro areas: 2020. Retrieved November 25, 2024, from https://cityobservatory.org/most_segregated2020/

Cutler, D., Glaeser, E., & Vigdor, J. (1999). The rise and decline of the American ghetto. Journal of Political Economy 107 (3): 455-506.

Deming, D. J. (2017). The growing importance of social skills in the labor market. The Quarterly Journal of Economics 132 (4): 1593-1640.

Dillon, F., Glaeser, E., & Kerr, W. (2025). Segregated workplaces and the careers of less-educated workers. Working Paper.

Duranton, G., & Puga, D. (2003). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), Handbook of regional and urban economics (vol. 4, pp. 2063–2117). Elsevier.

Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. Review of Economics and Statistics 97 (2): 498-520.

Kremer, M., & Maskin, E. (1996). Wage inequality and segregation by skill. NBER Working Paper No. 5718.

Lucas, R. E., Jr. (1988). On the mechanics of economic development. Journal of Monetary Economics 22 (1): 3-42.

Mas, A., & Moretti, E. (2009). Peers at work. American Economic Review 99 (1): 112-145.

Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. American Economic Review 94 (3): 656-690.

Murray, C. (2012). Coming apart: The state of white America, 1960–2010. Crown Forum.

Song, J., Price, D. J., Guvenen, F., Bloom, N., & von Wachter, T. (2019). Firming up inequality. The Quarterly Journal of Economics 134 (1): 1-50.

Taeuber, K.E., & Taeuber, A.F. (1965). Negroes in cities: Residential segregation and neighborhood change. Aldine Publishing Company.

US Census Bureau. 2000–2020. Longitudinal Employer-Households Dynamics Snapshot Release S2021 (LEHD), 2000-2020. Washington, DC: U.S. Census Bureau, Center for Economic Studies; distributed by Research Data Centers. https://lehd.ces.census.gov/data/lehd-snapshot-doc/latest/sections/introduction.html (accessed Jan. 9, 2025).

Appendix

AI. Segregation at different multi-unit geographic levels

While our focus is on worker segregation at the establishment level (defined as firm-blocks), workers might also benefit from working near other workers in the same firm in nearby establishments. To account for this, Appendix Figure 1 groups workers by firm and Core-Based Statistical Area (CBSA), and also by firm and state, to measure how segregation has changed across firms and broader geographic areas over time. This analysis also no longer relies on the establishment locations imputations in the LEHD's Job History File for workers of multi-unit firms. The new series diverge somewhat during the first ten years, with the CBSA- and state-based definitions showing modest declines, but thereafter the series increase in parallel.

Appendix Figure 1: Dissimilarity indices at multi-unit geographic scales, 2000-2020

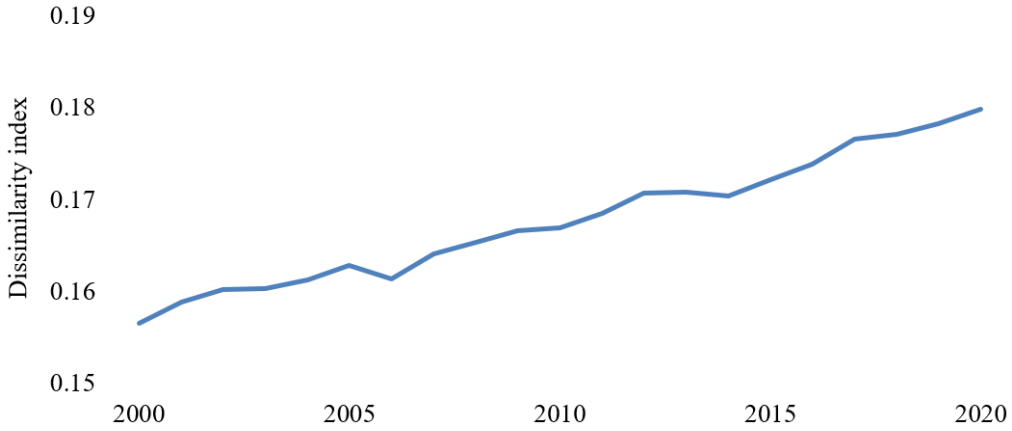
Notes: Figure plots changes in workplace isolation from their starting levels in 2000. Table 1 describes the underlying sample. The added series consider firm-CBSA or firm-state as the unit of observation for calculating worker segregation in multi-unit firms. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898, FY25-2766-R12120.

AII. Segregation across establishments within multi-unit firms

Segregation has also increased within multi-unit firms, as shown in Appendix Figure 2. Due to firm specialization, within-firm dissimilarity is far lower than across-firm dissimilarity. For example, a supermarket chain is unlikely to hire many college workers at any establishment, so across all establishments the measured within-firm dissimilarity will be low. This within-firm

segregation has grown over the past 20 years, indicating that multi-unit firms appear increasingly likely to employ college and non-college workers in different establishments.

Appendix Figure 2: Within-firm workplace dissimilarity index, 2000-2020



Notes: Figure plots the within-firm workplace dissimilarity index from 2000 to 2020. The index measures the share of the non-college workforce (or, equivalently, the college-educated workforce) that would need to move establishments in order to produce a uniform education share across workplaces for a firm. Table 1 describes the underlying sample. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY25-P2766-R12120.

AIII. Adjusting for incomplete worker samples

The LEHD provides imputed values for person-level characteristics when those values are not observed. Some characteristics, like age or sex, are derived from several sources and only imputed in around 5% of workers. Educational attainment, however, comes from rarer person-level surveys like the 2000 Decennial Census and the American Community Surveys and thus must be imputed in about 80% of observations during 2000-2020. While the imputations by the Census Bureau are estimated using many pieces of data, the randomness induced in imputed person-level educational attainment will bias establishments with more imputed workers to look less segregated. Additionally, the method used by the Census Bureau to impute worker education changed in 2009, keeping imputations from being longitudinally consistent. Thus, our main metrics use only workers with observed educational values to calculate segregation indices.

Yet, we still confront some empirical challenges with non-imputed data. In 2000, the sample of workers with non-imputed data is effectively missing at random because the long-form

2000 Decennial Census has just surveyed one in six of the population. In our sample, 29.85% of workers have observed educations in 2000. Over time, this share declines because workers with education values leave the LEHD but new workers are less likely to have their education levels picked up by the American Community Survey, which samples only one in a hundred each year.

Our estimations of segregation, which only use workers with observed educational status in each year, could thus suffer from three biases. First, even in 2000, our starting sample of the total population can provide an upward bias on segregation by making it look like workers are in smaller, more segregated establishments (i.e., data coarseness tends to increase measures of segregation). Second, over time, the sample of non-imputed workers decreases in relative terms, potentially upward biasing our estimates of changes in segregation for similar reasons. Finally, as we get farther away from 2000, we are less likely to observe the education of younger workers in particular, which means we have an additional downward bias on college shares, as younger workers are more educated generally than older workers.

Starting with the 2000 level, we perform two exercises to estimate what the likely level of segregation would have been without imputation issues. We first measure segregation using a 12%, 16%, 20%, 24%, and 28% random sample of workers with observed education in 2000. We then fit a quadratic regression on the observed indexes calculated using these five samples plus our actual sample of 29.85% of workers. The minimum value of this quadratic function across the 0% to 100% range provides one approximation of what a full sample would estimate. These minimum values for each series are documented in the second row of Appendix Table 1. While a bit smaller, they are quite close to the observed value.

Appendix Table 1: Levels of workplace segregation in 2000 with adjustments for imputed educations

	Isolation index, non-college	Adjusted isolation index, non-college	Isolation index, college	Adjusted isolation index, college	Dissimilarity index
	(1)	(2)	(3)	(4)	(5)
Observed values	0.787	0.108	0.549	0.228	0.519
Quadratic regression	0.787	0.108	0.546	0.225	0.516
Simulated values	0.787	0.108	0.514	0.193	0.495

Notes: See Table 1. Observed values are indices using all non-imputed values in 2000. The quadratic regression approach first estimates indices in 2000 using the observed value and in five random samples drawing 12%-28% of total workers. The minimum value of the fitted line between 0% and 100% is documented. The simulated value approach estimates indices at 12% and 29.85%, and compares the change in the true index with the corresponding change in a simulated index between 12%, 29.85% and 100% of values using a simulated dataset. By comparing the change in indices using different shares of observed values, we estimate what segregation would be at 100% observation. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898, FY25-2766-R12120.

The third row of Appendix Table 1 reports a second technique. We create a simulated population where we know the full sample segregation value. By randomly selecting a sample of 12% and 29.85% of this simulated group, we can model the under-estimates compared to a 100% sample. The complete calculation measures the likely full sample value as the observed segregation with 29.85% of workers plus the change in simulated segregation between 100% of workers and 29.85% of workers, with the latter scaled by the ratio of the observed over simulated change in segregation between 29.85% of workers and 12% of workers. The estimates from the simulation are very similar for non-college workers and modestly lower for college-educated workers.

We next turn to potential adjustments to the time trend of segregation to account for the general dwindling of our non-imputed sample over time and its particular weakening for young workers. Our first step calculates our indices for young, middle-aged, and older workers using similarly sized distributions of workers over time (i.e., restricting earlier years of the sample to look like later ones). We then weight these age- and size-consistent samples by the share of workers that fall into each group:

$$Index_{t,20-64} = \frac{Workers_{t,20-29}}{Workers_{t,20-64}} * Index_{t,20-29} + \frac{Workers_{t,30-49}}{Workers_{t,20-64}} * Index_{t,30-49} + \frac{Workers_{t,50-64}}{Workers_{t,20-64}} * Index_{t,50-64}$$

By combining the change in segregation using this method with the level-adjusted starting point in 2000 using a quadratic regression, we can repeat our core tables with an estimate of how they would have looked without imputation shortfalls.

Appendix Table 2 shows the revised adjusted non-college isolation index increases by more than before, while the adjusted college isolation index decreases. The dissimilarity index still increases from 2000 through 2020, but only between 2010 and 2020 after a slight dip in the first half of the series. In short, the emphasized metrics for non-college workers are very secure against imputation issues, while the college worker indices are sensitive. Appendix Tables 3 and 4 continue with the split outs for non-college workers, following Tables 2 and 3 of the main text. These tables are remarkably similar, indicating the robustness of the levels and trends in this emphasized part of our work.

Appendix Table 2: Comparison of data to alternative indices using imputation adjustments

	Adjusted isolation index, non-college DATA	Adjusted isolation index, non-college ALTERNATIVE	Adjusted isolation index, college DATA	Adjusted isolation index, college ALTERNATIVE	Dissimilarity index DATA	Dissimilarity index ALTERNATIVE
	(1)	(2)	(3)	(4)	(5)	(6)
2000	0.108	0.108	0.228	0.225	0.519	0.516
2010	0.126	0.133	0.226	0.200	0.525	0.511
2020	0.147	0.169	0.240	0.188	0.550	0.527
Δ 2000-2020	0.039	0.061	0.012	-0.037	0.031	0.011

Notes: Columns 1, 3, and 5 show non-imputed indices from Table 1. Columns 2, 4, and 6 document estimated indices with adjustments for dropping non-imputed workers. These alternative indices set the level of workplace segregation in 2000 using a quadratic regression methodology (see Appendix Table 1) and calculate the change in workplace segregation using an age- and sample-size corrected index (see Appendix Formula 1). Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY24-P2766-R11646, R11898, FY25-P2766-R12120.

Appendix Table 3: Table 2 with alternative indices using imputation adjustments

	Adjusted isolation index, 2000	Dissimilarity index, 2000	Δ Adjusted isolation index, 2000-2020	Δ Dissimilarity index, 2000-2020
	(1)	(2)	(3)	(4)
Total	0.108	0.519	0.061	0.011
Men	0.185	0.520	0.070	0.023
Women	0.040	0.511	0.051	-0.002
White	0.125	0.519	0.064	0.006
Minority	0.089	0.502	0.050	0.034
Young (18-29)	0.151	0.514	0.094	0.017
Middle Aged (30-49)	0.114	0.516	0.030	0.018
Older (50-64)	0.098	0.516	0.088	-0.003
Northeast	0.114	0.513	0.049	0.004
Midwest	0.127	0.500	0.053	0.001
South	0.123	0.526	0.056	0.000
West	0.113	0.519	0.069	0.024

Notes: See Table 2 and Appendix Table 2. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY25-P2766-R12120.

Appendix Table 4: Table 3 with alternative indices using imputation adjustments

	Adjusted isolation index, 2000	Dissimilarity index, 2000	Δ Adjusted isolation index, 2000-2020	Δ Dissimilarity index, 2000-2020
	(1)	(2)	(3)	(4)
Total	0.108	0.519	0.061	0.011
Goods-producing firm	0.192	0.479	0.068	0.010
Services-producing firm	0.086	0.500	0.064	0.014
Single-unit firm	0.142	0.593	0.063	-0.006
Multi-unit firm	0.099	0.451	0.058	0.034
Small establishment	0.185	0.619	0.070	-0.011
Medium establishment	0.103	0.461	0.056	0.015
Large establishment	-0.024	0.366	0.033	0.015
New firm	0.159	0.618	0.088	0.040
Existing firm	0.114	0.503	0.059	0.015

Notes: See Table 3 and Appendix Table 2. Disclosure conducted under FSRDC Project Number 2766. CBDRB-FY25-P2766-R12120.