Individual and Social Genomic Contributions to Educational and Neighborhood Attainments: Geography, Selection, and Stratification in the United States

Thomas Laidley, Justin Vinneau, Jason D. Boardman

University of Colorado Boulder

Abstract: Research on neighborhood effects draws suggestive links between local spatial environments and a range of social, economic, and public health outcomes. Here, we consider the potential role of genetics in the geography of social stratification in the United States using genomic data from the National Longitudinal Study of Adolescent to Adult Health. We find that those with genotypes related to higher educational attainment sort into neighborhoods that are better educated and have higher population densities, both descriptively and using formal school and sibling fixed-effects models. We identify four mechanisms through which this geographic sorting on genetic endowment can magnify social stratification: assortative mating, social-genetic effects, gene-by-environment interactions, and gene–by–social-genetic interactions. We examine the presence of the latter three in our data, finding provisional yet suggestive evidence for social-genetic effects that putatively amount to about one-third of the influence of one's own genomic profile. We find no evidence, however, for the presence of interactions between environments and individual genetic background. Collectively, these findings highlight the potential for geographic sorting on genotype to emerge both as a key methodological concern in population genetics and social science research and also a potentially overlooked dimension of social stratification worthy of future study.

Keywords: neighborhood effects; social genome; spatial stratification; neighborhood attainment

A large body of social science scholarship has persuasively emphasized the importance of geography in unpacking social, economic, and health outcomes. Cities, neighborhoods, and other local environments are thus conceptually understood not merely as passive spatial "containers" of granular behavior but influential forces in their own right that affect trajectories over and above individual- or family-level factors. From the work of Wilson (1987), who documented the social and economic dislocation of black Americans induced by the persistent spatial concentration of urban poverty wrought by deindustrialization, to Chetty and colleagues (2014; 2018), who have animated a growing body of literature on the geography of mobility and opportunity in the United States (Aghion et al. 2019; Berger and Engzell 2019; Sampson 2019), researchers have provided much suggestive evidence on the importance of space and place in untangling the contours of stratification and inequality. The recent development of genomic indicators of phenotypes related to social outcomes produced by genome-wide association studies (GWAS) and their integration with survey data offers a novel opportunity to explore whether

the geographic distribution of specific genotypes may play a heretofore largely unexamined role in producing (and reproducing) structural disadvantage.

Here, we focus on the phenotype of educational attainment, a straightforward and powerful factor in determining economic mobility and labor market success (Hout 2012; Torche 2011). Using survey data from the National Longitudinal Study of Adolescent to Adult Health (Add Health) along with a newly developed polygenic scores (PGSs) for educational attainment (Lee et al. 2018), we first descriptively document whether differential neighborhood attainment in adulthood and long-distance geographic migration is associated with educational PGS. We examine whether respondents with higher PGSs for educational attainment were more geographically mobile, made more long-distance (>25 miles) moves between adolescence and young adulthood, and moved into more advantaged neighborhoods (or remained in environments that realized outsized long-run secular gains compared to their peers[i.e., those with a higher proportion of college-educated residents, higher home values, etc.]). After finding suggestive descriptive evidence of sorting and variation in regional mobility on genotype, we turn to more formal analyses, using school fixed effects and sibling-difference models to illustrate how genetic background predicts neighborhood characteristics in adulthood. Finally, with evidence of residential sorting on genotype over time, we consider whether the systematic migration of individuals with higher PGS for educational attainment into more advantaged environments could magnify social and economic stratification over and above the direct influence of neighborhood attainment through four hypothesized mechanisms: (1) assortative mating, (2) "social-genetic effects" (i.e., the influence of the genome of one's environmental peers over and above one's own genetic background), (3) gene-by-environment interactions (GxE; the moderation of genetic associations by the social environment), and (4) and "social epistasis," or the moderation of the influence of one's own genome by the "social genome" (gene–by–social-genetic interaction [SGxG]). Because testing for marriage homogamy is not feasible with the data in Add Health, we focus on the latter three phenomena in our analyses. We also discuss the potential methodological implications of selection processes, including on the results we obtain here.

This article makes important contributions to the work on spatial inequality and putative neighborhood effects in several ways. First, we descriptively show how neighborhood attainments in adulthood correlate with alleles related to education, whereas more formal sibling fixed-effects models suggest that this relationship is plausibly causal and not merely reflective of social factors like parental education or other resources (e.g., neighborhood amenities) common within families. Although this apparently active selection on genetic endowment has methodological implications for research examining neighborhood effects and gene-by-environment interactions—as it demonstrates that social environments are not strictly exogenous to genetics—we focus on what influence this finding may have on broader trends in inequality and mobility. We find suggestive associations between the genetic environment (i.e., the average PGS of one's neighborhood peers) at the neighborhood level during childhood and educational attainment in adulthood even among respondents in common school environments and net a suite of social covariates. Taken together, these findings suggest that integrating genetics into

social science research has the potential to add to our understanding of phenomena like neighborhood effects, attainments, and stratification insofar as it is both socially and individually relevant for environmental selection and educational credentials.

The Genetics of Education and Selection on Genome

PGS are quantitative indicators of genetic influence[1] and are crafted by drawing associations between the variation in specific single-nucleotide polymorphisms (SNPs)—or variations of specific alleles in the genome across individuals—and traits (or phenotypes) of interest. The relative influence of specific SNPs varies by the strength of their association with the phenotype and is weighted accordingly and summed to produce an overall standardized score that reflects genetic disposition to exhibit a given trait. Associations between genetic background and social environments (i.e., gene-environment correlations, or rGE) may be the product of active, passive, or evocative selection mechanisms. In active selection, individuals seek out environments that better comport with their behavioral predilections, which are themselves partly genetically patterned. For example, an individual with a PGS that is linked to an increased likelihood of obtaining a college degree may actively select into environments with more educational resources or cultural amenities. In passive selection, environments are merely inherited, as they would be for the children of parents who perhaps themselves actively selected into places based on their genetic profiles; insofar as children share their genetic background with parents, such correlations would persist intergenerationally. Evocative selection is a process whereby genes evoke a selection response indirectly. For example, a student may exhibit behaviors or skills in classes that are linked to his or her cumulative genotypes, and teachers may respond to these behaviors by placing the student in a context that further fosters growth and development. In this case, genotype is linked to an outcome because it evokes an environmental response that promotes the outcome.

To tease out whether PGS indicators reflect causal, influential genetic variation in the first place rather than merely reflecting signatures of social privilege or cultural differences, some work has used a sibling-difference analytic approach that exploits within-family genetic "lotteries" and zeroes out common environmental, ancestral, and indirect genotypic influences (i.e., parents' genotype indirectly influencing outcomes through "genetic nurture" [Kong et al. 2018], which has been conceptualized as a social-genetic effect in the literature [Liu 2018]). For instance, recent work shows that siblings with higher PGSs for education realize significantly higher degree attainment and greater intergenerational mobility using family background as a fixed effect (Belsky et al. 2018). Indeed, recent GWAS that produce one of the newest vintages of PGS for educational attainment (which our analyses are based on) implicate a number of SNPs that are involved in neurological functioning and point to plausibly biophysical pathways between the genome and outcomes that suggest cognitive ability as a mechanism (Lee et al. 2018).

Still, the residual gap between the predictive power of PGSs with respect to cognitive performance and educational attainment—along with the lack of research that directly examines putatively biological pathways between genetic variation

and phenotypes—preclude any confident claims as to what extent scores reflect ability. Even if education PGSs are not culturally patterned or population structure is convincingly accounted for in a research design, it is hardly clear with the current dearth of biological evidence that it reflects cognitive performance rather than other behaviors or characteristics that may be advantageous in higher education or the labor market but meaningless in any substantive sense (e.g., physical features or appearance). Moreover, genetic variants related to education have been associated with outcomes that suggestively operate through indirect channels over and above the phenotype (i.e., degree acquisition), which further complicates drawing causal pathways between the genome and outcomes. For instance, recent work associates alleles for educational attainment with wealth accumulation through a process not entirely mediated by college credentials, which suggests more nebulous pathways of risk preference and other behavioral complexities that are not reducible to "ability" or cognition per se(Barth, Papageorge, and Thom 2018). For this reason, research that examines the influence of PGS related to a complex social phenotype like educational attainment must remain largely circumspect about whether, say, PGS predicts college completion through a channel like cognitive skill as opposed to more nebulous linkages or some combination thereof.

Spatial Sorting, Neighborhood Effects and Attainments, and Mechanisms of Stratification

Although estimating causal neighborhood or peer effects is notoriously fraught due to concerns of self-selection (Angrist 2014; Feld and Zölitz 2017; Graham 2018), and identifying concrete mechanisms of action is difficult even with more robust analytic approaches (Minh et al. 2017), recent research that uses a variety of methodological tools (e.g., instrumental variable approaches or natural experiments) to address bias on unobservables seems to generally suggest the importance of the social environment in contributing to outcomes over and above factors at the individual and family levels. For instance, recent work in the United States has found measurable contributions of school and neighborhood environment on postsecondary school performance (Galster et al. 2016) and educational attainment and wages in adulthood (Altonji and Mansfield 2018; Chetty and Hendren 2018), whereas similar research in the Netherlands documents a significant association between neighborhood deprivation and adult earnings (van Ham, Boschman, and Vogel2018).

Still other research in this vein has found that the effect of neighborhood disadvantage on outcomes like school performance or teenage pregnancy hinges on the recency and timing of exposure in addition to the cumulative total, which stresses the likelihood that individuals may be more sensitive to their surrounding environment during specific stages of the life course but particularly adolescence (Hicks et al. 2018; Kleinepier and van Ham 2018; Levy 2019; Levy, Owens, and Sampson 2019; Wodtke 2013; Wodtke, Elwert, and Harding 2016). Results on the links between neighborhoods and other social environments and health outcomes like body mass index (BMI) are somewhat equivocal (Arcaya et al. 2016)—largely owing to the

methodological challenges of working with observational epidemiological data—yet there is some evidence of general proximity effects, whereby obesity spatially diffuses on a regional scale over time (Agovino, Crociata, and Sacco 2019) and is influenced by poverty and demographic composition at the local levels (Yang and South 2018).

Work on neighborhood attainment in the United States—how, where, and why individuals sort into and out of more or less advantaged environments over time or how neighborhood conditions change around "stayers" who remain in situ—has largely focused on the ethnic and racial dimensions of neighborhood composition and sorting due to the historical pervasiveness and severity of residential segregation. In their study of the legacy effects of the Great Migration, Leibbrand et al. (2019) found compelling evidence that second-generation black families who migrated north were more successful at translating socioeconomic status (SES) into improved neighborhood conditions compared to those who stayed in the South. Research examining more recent birth cohorts, however, has found that African Americans are less likely to convert baseline neighborhood or family-level advantages into moves up the residential opportunity matrix compared to whites (Brazil and Clark 2017; South et al. 2016) or second-generation immigrants (Tran 2019). Other research that decomposed differences in neighborhood attainments between whites and blacks found that in situ changes in neighborhood conditions largely explain the growth in these gaps along the life course (Huang, South, and Spring 2017), though long-distance (i.e., inter- vs. intracounty) moves do tend to be associated with marked improvement in local environments across subgroups (Sampson and Sharkey 2008; South et al. 2016). Other research using instrumental variable and formal mediation techniques to identify plausibly causal through-lines between residential segregation and attainment finds suggestive evidence of neighborhood effects acting as the nexus between initial conditions and residential context in adulthood (Pais 2017).

Research examining the spatial distribution of PGSs related to outcomes of interest to social scientists (e.g., educational attainment, BMI, depression, etc.) rather than minor ancestral differences that are evident across the genome and the methodological bias they can induce is more limited because of the relatively recent integration of genomic indicators into survey data. In Great Britain, researchers using data from the UK Biobank project found evidence that PGS related to educational attainment exhibits a significant level of spatial clustering and that this distribution is patterned based on SES in the expected fashion, with lower-SES regions containing populations with a greater frequency of attainment-decreasing alleles (Abdellaoui et al. 2018). In the United States, recent work examining clustering at the state level using Health and Retirement Study (HRS) data found relatively modest levels of homophily of PGS for educational attainment (Rehkopf, Domingue, and Cullen 2016), which also appears to remain fairly stable across the life course for these cohorts (Domingue et al. 2018b). Contemporaneous work undertaken independently from ours, however, has shown that risk alleles related to lower educational attainment and lower age at first birth are associated with neighborhood deprivation indices and downward intergenerational mobility among respondents in Add Health (Belsky et al. 2019). This suggests that modest levels of state-level

clustering or across-state variation in average PGS may belie more uneven spatial distribution patterns and genetic assortment at smaller levels of geography, such as the county or neighborhood.

Aside from the possible direct benefits of converting genetic endowment into better neighborhood conditions in adulthood, the geographic clustering of a genetic predisposition for education may further impact social and spatial stratification outcomes if mechanisms exist that act to increase baseline gaps based on the interplay between social environments, individual genetic background, and the genetic background of proximate individuals. Perhaps the most straightforward implication of spatial sorting by genetic background with respect to social stratification is assortative mating, whereby populations with higher PGS for attainment sort into environments where they are more likely to form families with like individuals and vice versa. Although teasing out the pathway between spatial genetic homophily, assortative mating, and intergenerational outcomes is complicated by data availability (and, indeed, impractical with the dataset we use here), there is some suggestive evidence of general genetic assortative mating in the United States (Domingue et al. 2014) and the United Kingdom (Hugh-Jones et al. 2016). Still, the available evidence suggests that genetic homogamy is rather trivial compared to that based on social factors like educational attainment (Conley et al. 2016a), and the interpretation of results is complicated by considerations of how much it is attributable merely to coupling on ancestry and ethnicity (Abdellaoui, Verweij, and Zietsch 2014).

It is also possible that the "social genome" (i.e., the average genetic endowments of one's peers in a local environment) can confer additional advantages on individuals, which would further magnify stratification patterned on geography. Although the application of the peer effects conceptual framework to population genetics research is relatively novel, the limited extant work does suggest that the social-genetic environment—that is, the average PGS for a given trait among one's proximate peers—may have an influence over and above one's own genetic endowment. Using Add Health data, research has found average educational attainment PGS among alters from both schoolmate and friendship networks to be significantly associated with educational outcomes, net the ego's own PGS (Domingue et al. 2018a). Similar work finds that state-level mean PGS for education predicts attainment net the individual's score in HRS data (Domingue et al. 2018b). Other recent research deploys quasiexperimental analytic strategies to draw a plausibly causal through-line between the genetic propensity to smoke among school alters, peer smoking behavior, and ego smoking behavior (Sotoudeh, Conley, and Harris 2017). In a similar line of work, researchers have also examined the influence of the parents' genomes on educational outcomes net the children's own endowment, whereby alleles that are not transmitted intergenerationally nevertheless have a putative effect on outcomes through environmental channels, or "genetic nurturance" (Belsky et al. 2018; Kong et al. 2018; Liu 2018).

Another consideration is whether there is a moderation of the effect of one's own genome based on that of one's peers—a SGxG interaction, or social epistasis—with similarly suggestive but provisional evidence that the genome of schoolmates and friends significantly affects the magnitude of the effect of one's own genetic endowment on completed education (Domingue et al. 2018a). Finally, there is

the possibility of GxE interactions, whereby social environments—for example, the proportion of the neighborhood that is college educated or living in poverty—moderate the influence of one's own genome on outcomes of interest. Although GxE (or SGxG) analyses are subject to the same methodological pitfalls involving possible confounding on selection (and have been persuasively argued to be altogether misspecified in a great deal of past work [Keller 2014]), recent research suggests that school environments significantly moderate the association between genetic endowments and college completion in the United States (Trejo et al. 2018). Other research examines how genetic penetrance—or the power of a given PGS to predict an associated phenotype—changes across birth cohorts, implicitly modeling associated unique differences in historical context as environmental conditions. Recent work in this vein finds that education PGS has become less predictive with respect to actual attainment in more contemporary cohorts, presumably because of easier access to education over time (Conley et al. 2016b).

Materials and Methods

Data

Add Health is a nationally representative study of children and adolescents who were in grades 7 to 12 in the 1994 to 1995 school year and consists of about 12,000 respondents from more than 130 schools in the core study, or more than 20,000 including supplemental oversamples (e.g., twins, siblings, etc.). Subsequent waves of data collection consisted of an immediate follow-up in 1996 (wave 2), another from 2001 to 2002 when children were transitioning to adulthood (wave 3), and another from 2007 to 2008 when they were approximately 24 to 32 years of age (wave 4).

Measures

Neighborhood characteristics and average polygenic scores. Though respondents in Add Health are not assigned geocodes that indicate specific location, supplementary indicators based on census data (1990 for the first wave, 2005–2009 American Community Survey data for the fourth) are available at the county, tract, and block group level for the first two waves, and at the tract level for the latter two. For our descriptive analyses, we use indicators of population density (thousands of persons per square mile), proportion of the population that is older than 25 years of age with a college degree, median home value, and median household income at the census tract level. For our main predictors of interest in the social-genetics estimations, we first calculate leave-out averages for educational attainment PGS at the census tract level. That is, we calculate the average PGS of a given respondents' neighborhood alters within a common census tract in the first wave (see Figure 1).

Though specific identifiers are unavailable, Add Health provides grouping codes that can be used to place respondents in common neighborhood environments. Though the number of respondents used to calculate these average PGS figures at the tract level are fairly robust (median = 26; mean = 37.67), there are some cases
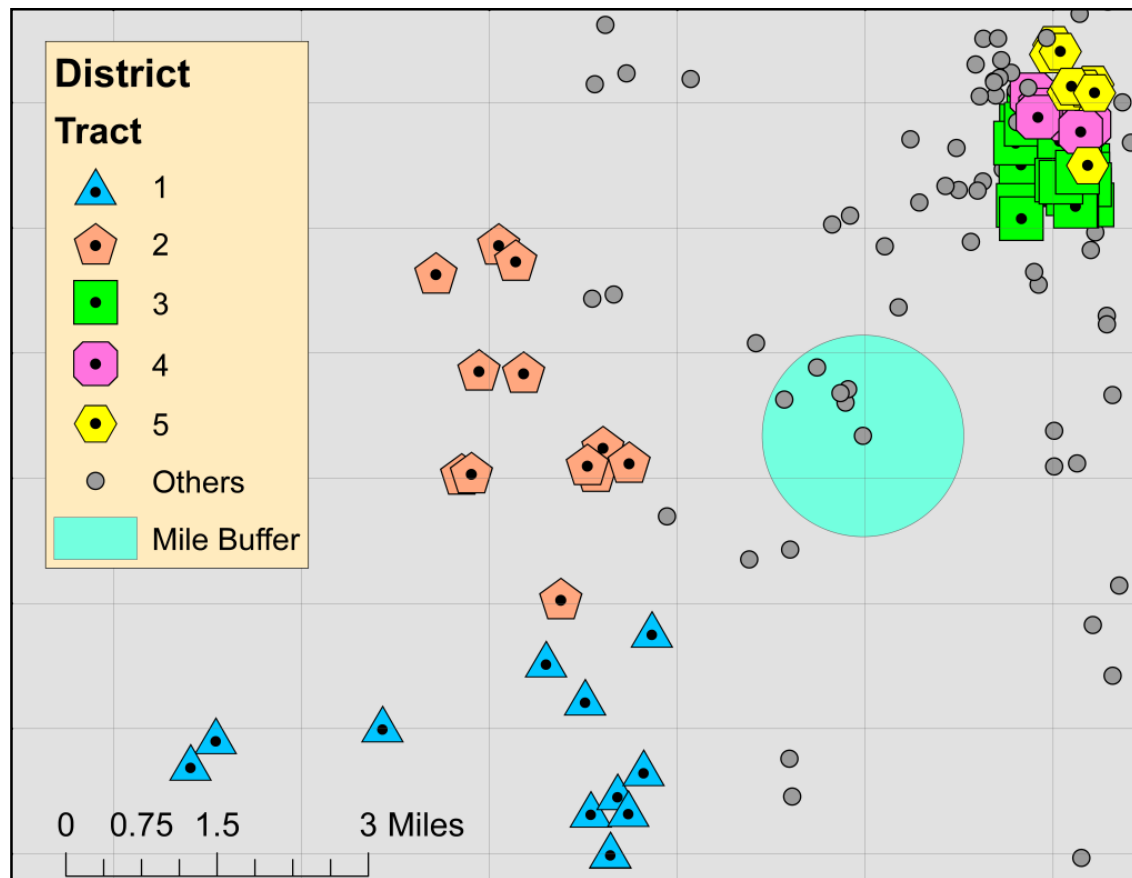
Figure 1: The spatial dispersion of individuals within a common school district and their association with different census tracts within that common environment. Spatial data were jittered/dispersed randomly to obscure the relative location of individuals, and only a subset of the total is shown. The five tracts pictured have the most affiliated respondents in this given school district; others are shown as gray dots. The one-mile buffer node was chosen randomly for illustration and has no real-world significance (i.e., it does not signify the school or other municipal feature).

where relatively few numbers of alters are used to construct means. We undertake sensitivity analyses that deploy different thresholds of the number of neighborhood alters used to construct tract-level measures and discuss those in the results. In an effort to supersede tract boundaries, which often correspond to geographic or infrastructural features but are ultimately arbitrary delineations, we also used spatial data that assigned dummy coordinates to respondents within common communities in the first wave. From these raw spatial data, we constructed a distance-weighted, leave-out average PGS based on peers within a one-mile radius. Because many respondents in Add Health live in environments that are more rural in character (or more suburban within urban macroenvironments), we obtained fewer valid observations for this than tract-level measures, which capture anybody

within a common census boundary. Figure 1 graphically illustrates an example of this dispersion.

Educational and geographic mobility outcomes. We used two indicators of educational attainment in adulthood, derived from the fourth wave, when the youngest cohort was about 24 years old. The first is whether the respondent had attained a four-year bachelor's degree. The second is total years of reported education. For our descriptive and school/family fixed-effects estimates, we used indicators of the total distance in miles between waves and primarily concentrate on the total distance between locations at waves 1 and 4. From this, we also calculated a binary indicator of whether the respondent moved more than 25 miles, which we use to proxy interregional moves. We also include relevant covariates, including age, sex, and cognitive performance at the first wave (as captured by Peabody Picture Vocabulary Tests). For mixed-effects models predicting educational attainment that do not account for family environment, we additionally control for parental education and the proportion of the wave 1 tract that was college educated to capture neighborhood conditions. Descriptive statistics for measures described here are available in Table A1 of the online supplement.

Analytic Approach

For descriptive analyses, we took average neighborhood and geographic mobility figures within education PGS quintiles (see Table A2 of the online supplement for means and standard deviations by quintile). We then employ a more formal school and sibling fixed-effects approach that examines whether individual genetic background is associated with neighborhood attainment outcomes over and above the associated social factors like degree attainment or parental SES. First, we estimated neighborhood attainments in adulthood using the school environment as a fixed effect and control for other individual factors like parental education and cognitive performance. We then separately analyzed neighborhood attainments among respondents of any relation in common household environments and, more conservatively, among full siblings and dizygotic twins. By modeling the home and family environment as a fixed effect, we more powerfully account for resources that vary between and not within households, like wealth, parental education, and neighborhood. In the school and family fixed-effects estimates, we control for sex, age, cognitive performance in the first wave, wave 4 state fixed effects, and completed education in the fourth wave. Finally, in our models examining social-genetic effects and interactions, our preferred estimates used a mixed-effects approach where schools are used as the grouping variable and education outcomes are predicted with tract-level and spatially rendered PGS, net age, sex, cognitive performance in wave 1, proportion of wave 1 census tract college educated, and parental education. Because we found tract-level college education to evince strong associations with our outcomes of interest, whereas other variables like proportion in poverty or racial or ethnic composition add nothing significant to our predictions, we parsimoniously exclude other environmental indicators. As part of the wave 4 data collection, saliva samples were obtained from consenting participants (96 percent of wave 4 respondents), and a total of 12,200 agreed to have their sam-

ples archived for future projects. Genotyping was performed using the Illumina HumanOmni1-Quad BeadChip and the Illumina Omni2.5-Quad BeadChip. After quality control procedures described in great detail elsewhere (Highland et al. 2018), genome-wide data are available for 9,974 individuals.

Polygenic Scores for Education

Our analyses use the PGS for education that is made available to Add Health researchers (Braudt and Harris 2018). Briefly, a PGS is simply the product of the effect size for a particular allele and an individual's genotype at that loci, summed across the entire genome. This process is implemented using the PRSice wrapper for *R* that invokes the Plink software. Detailed information for this method is described by Chang et al. (2015).

## Population Stratification and Ancestry

Much of the work on the geography of genetics and its spatial distribution focuses on ancestral differences insofar as they constitute fundamental methodological concerns both for researchers who construct PGS indicators and those that apply them empirically. Ancestral differences that are patterned by geography can confound such associations insofar as they may dovetail with very small differences in allele frequencies and cultural variation in phenotypes and induce correlations that are not truly causal in nature. In the paradigmatic "chopsticks gene" example (Hamer and Sirota 2000), researchers drawing associations between genetic variants and the use of chopsticks would pick up associations that reflected East Asian ancestral markers rather than alleles that substantively affected utensil use, as it is a socially and culturally but not biologically patterned behavior. Indeed, early work in this vein documented a rather robust association between genetic variation and geographic location on global (Wang, Zöllner, and Rosenberg 2012) and continental scales (Novembre et al. 2008). More recent research that benefits from comparatively advanced computational power and refined methodology goes even further and documents subtle but nontrivial differences in genetic variation reflecting ancestral differences even within ostensibly homogenous populations, including within the United Kingdom (Haworth et al. 2019), Finland (Kerminen et al. 2017), and Italy (Sazzini et al. 2016). These subnational regional differences in the distribution of alleles demonstrably co-occur with variation in phenotypes; thus, including markers of ancestry in contemporary population genetics research is a crucial first step in addressing (but not definitively solving [Berg et al. 2019]) methodological bias induced by even very fine–grained population stratification. In keeping with best practices, we relegate our analyses to those with primarily European ancestry, as PGSs are poorer predictors of phenotypes for other groups, and control for the top 20 principal components related to ancestry in all models except for those using full siblings and dizygotic twins.

Results

Descriptive Findings on Geographic Mobility and Neighborhood Attainments

We first descriptively analyzed how PGS for educational attainment relates to neighborhood conditions and geographic migration from adolescence to adulthood in Add Health. Our descriptive analyses are based on a subsample of unrelated, European-ancestry respondents who were successfully genotyped and had neighborhood information available for both the first and fourth waves in 1994 to 1995 and 2007 to 2008, respectively ($n = 4{,}013$). We base our findings on respondents with European ancestry—which is customary in the genomics literature—because GWAS are overwhelmingly composed of these respondents, and thus PGS derived from them are demonstrably poorer at predicting outcomes using more diverse populations (Lee et al. 2018:1116). We unpack the implications of this crucial limitation further in our discussion. We construct quintile bins of PGS for educational attainment for these respondents and report means and 95 percent confidence intervals (CIs) for tract-level educational attainment (proportion of those 25 years old or older with college degrees), median owner-occupied home values, household income, and population density at waves 1 and 4, when the children were adolescents and young working-age adults, respectively (see Figure 2 for a visualization and Table A2 in the online supplement for corresponding values).

We find significant baseline differences between the highest and lowest PGS quintiles in neighborhood education, household income, and home values. For instance, the proportion of the home census tract with a college degree in adolescence was about 30 percent higher for the top quintile (mean = 26.99; 95 percent CI: 25.90–28.0) compared to the bottom (mean = 20.60; 95 percent CI: 19.8–21.40) in the first wave. The only neighborhood characteristic that exhibited no statistically significant baseline differences between quintiles was population density (see Table A2 in the online supplement for one-way analysis of variance significance levels). However, gaps between the lowest and highest PGS quintiles in tract educated, median home value, and particularly population density grew significantly as respondents aged into adulthood. Although those in the top PGS quintile lived in tracts that were about 20 percent more densely populated at the first wave, by the fourth wave, it was more than double that of their peers in the lowest quintile. We observed similar trends with respect to long-term geographic mobility as measured by total miles between first and fourth wave locations and the proportion who moved more than 25 miles from their childhood home, which we use to proxy interregional as opposed to within-metropolitan relocations. Respondents in the top quintile, on average, migrated 257 miles between adolescence and adulthood (95 percent CI: 220.12–294.43), compared to 132 miles for their peers in the lowest quintile (95 percent CI: 109.50–156.50), and were about 60 percent more likely to migrate more than 25 miles over the same timeframe.
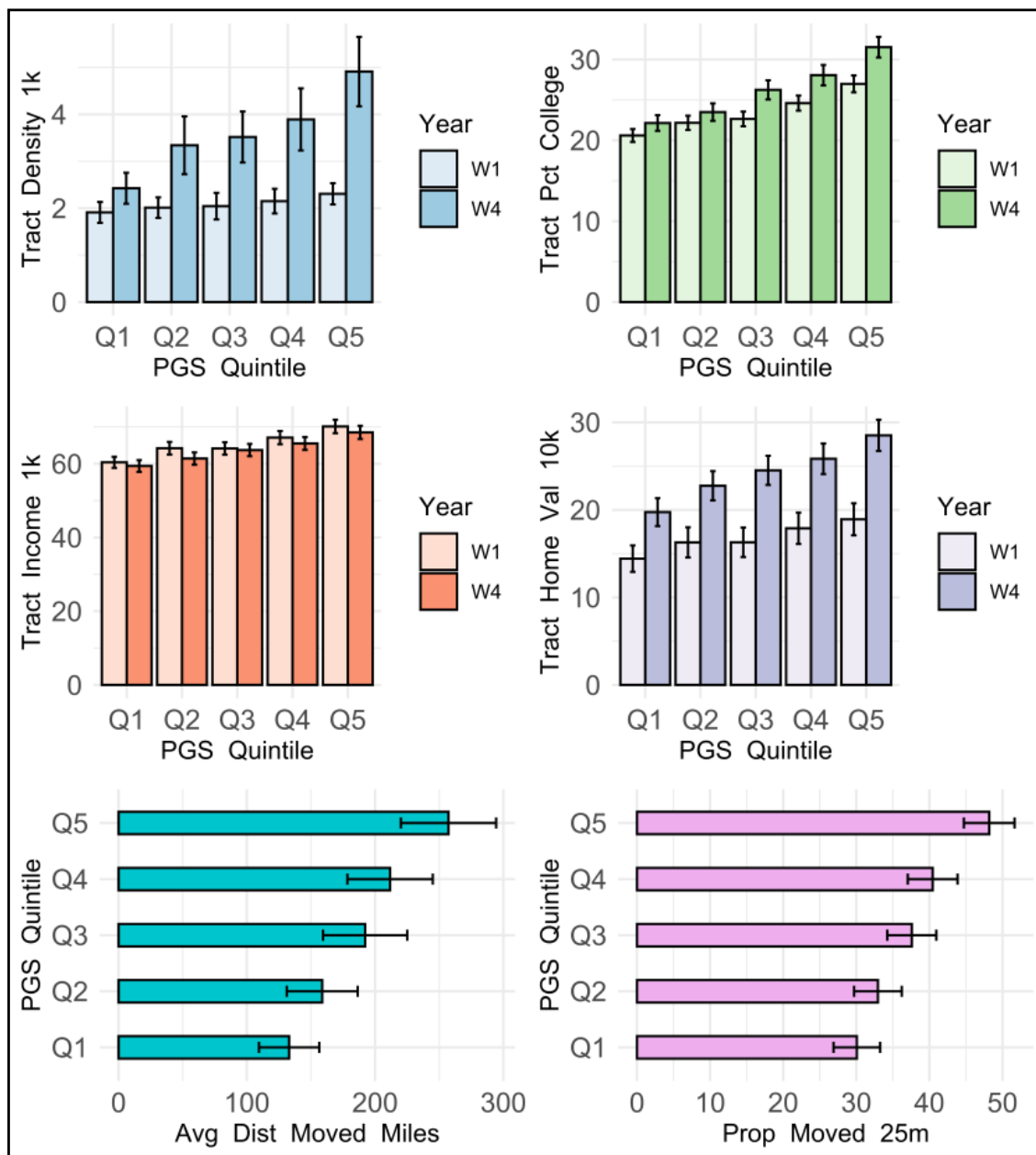
Figure 2: Means and standard errors for neighborhood characteristics in waves 1 and 4, distance between wave 1 and wave 4 locations, and proportion who migrated more than 25 miles by education PGS quintile. 1k, 1,000; Avg, average; Dist, distance; m, miles; Pct, percentage; Prop, proportion; Q1, quintile 1; Q2, quintile 2; Q3, quintile 3; Q4, quintile 4; Q5, quintile 5; Val, value; W1, wave 1; W2, wave 2; W3, wave 3; W4, wave 4.

School Fixed Effects and Sibling-Difference Models

Although our descriptive estimates suggest geographic sorting on genotype, we seek to test whether individual PGS predicts neighborhood attainment and mobility using a more formal approach in which we can account for covariates that relate to both genetic endowment and the school and home environment. Using school as a fixed effect and controlling for educational attainment in adulthood among other factors, we find that the PGS for education significantly predicts the population density, the proportion of the population with a college education, and median home value in neighborhoods in wave 4, along with the likelihood of having migrated more than 25 miles between the first and fourth waves (see Table 1 for results and Figure 3 for visualized regression lines). Although the magnitude of the effects is modest compared to actual educational credentials—even a two–standard deviation increase in educational PGS would be one-third of the effect size of a degree on tract-level proportion of those who were college educated in wave 4—genetic background appears to influence geography over and above a litany of other factors.

We next performed analyses using household fixed effects and sibling-difference models, wherein we exploit the random variation in alleles related to education within families to draw a more convincing through-line to outcomes of interest. By looking at full siblings and dizygotic twins, who share a part of their genome—on average about half—but have the same family background and neighborhood environments, we resolve questions of confounding on ancestry as a matter of course and zero out the influence of traits like parental education, parental genetic endowment and the indirect environmental pathways through which it operates, neighborhoods, and other contextual conditions. In these models, the PGS for education significantly predicts a standardized natural log of density and proportion of those who are college educated in wave 4 census tracts (see Tables 2 and 3 for results and Figure 4 for visualized regression lines) but not other neighborhood conditions. Estimates were rather similar using either household or the more conservative sibling fixed effects. Notably, educational PGS within households and sibling dyads significantly predicted wave 4 tract population density, whereas having a college degree did not. A college degree significantly predicted the proportion of wave 4 tract college educated only in the household fixed-effects model (see column 3 of Table 2).

Social-Genetic Associations and Interactions

To explore potential social-genetic associations, we predict both degree attainment and years of education in the fourth wave using measures of tract-level average education PGS and a distance-weighted average within a one-mile radius centered on each respondent in the first wave (see Materials and Methods). These measures capture the average PGS of Add Health peer respondents in the same neighborhoods (census tracts) or spatial buffer zones, not including the individual for whom the average is being obtained (i.e., leave-out averages). Our preferred estimates (see Tables 4 and 5) are based on a mixed-effects approach that allows random slopes
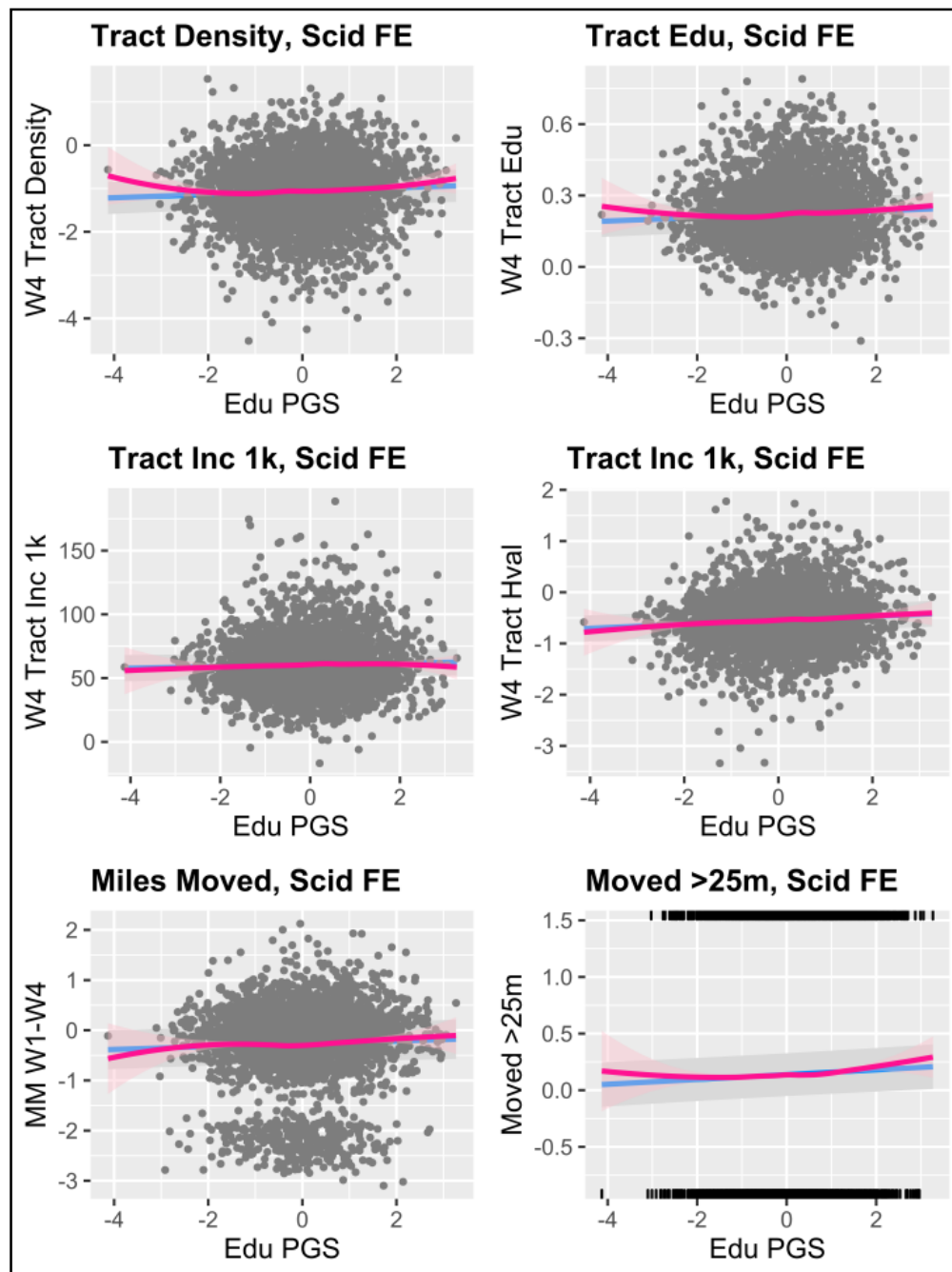
Figure 3: Visualized regression lines for individual educational PGS predicting neighborhood conditions in wave 4 and distance traveled between waves 1 and 4 in school fixed-effects models. Regression lines (blue for linear, pink using Loess smoothing) are consistent with the estimates contained in Table 1 and produced using the visreg package in R (see Breheny and Burchett [2017]). Edu, educated; FE, fixed effects; Inc, income; MM, miles moved; Scid, school ID.

Table 1: Wave 4 neighborhood attainment outcomes; unrelated, European subsample, wave 1 school fixed effects.

	Density	Educated	Income	Home Values	Miles Moved	Moved $> 25M$
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.47	0.15	17.81	0.10	0.79	1.11[†]
	(0.63)	(0.11)	(17.44)	(0.43)	(0.66)	(0.33)
Education PGS	0.04[†]	0.01[†]	0.64	0.04[†]	0.03	0.02[†]
	(0.01)	(< 0.01)	(0.40)	(0.01)	(0.02)	(0.01)
Degree Wave 4	0.14[†]	0.06[†]	5.07[†]	0.21[†]	0.22[†]	0.19[†]
	(0.03)	(0.01)	(0.91)	(0.02)	(0.03)	(0.02)
Cognitive Score	0.06[†]	0.02[†]	0.83	0.05[†]	0.05[†]	0.03[†]
	(0.02)	(< 0.01)	(0.53)	(0.01)	(0.02)	(0.01)
Age	-0.01	< 0.01	0.69*	< 0.01	< 0.01	-0.01
	(0.01)	(0.00)	(0.30)	(0.01)	(0.01)	(0.01)
Male	0.06*	-0.00	-1.83*	-0.02	-0.04	< 0.01
	(0.03)	(0.00)	(0.76)	(0.02)	(0.03)	(0.01)
Parent College	0.10[†]	0.03[†]	1.70	0.06[†]	0.04	0.07[†]
	(0.03)	(0.01)	(0.96)	(0.02)	(0.04)	(0.02)
Tract College	0.05*	0.03[†]	3.55[†]	0.10[†]	< 0.01	0.01
	(0.02)	(0.00)	(0.65)	(0.02)	(0.02)	(0.01)
Model	FE	FE	FE	FE	FE	FELogit
R^2	0.47	0.41	0.31	0.68	0.27	
Adjusted R^2	0.43	0.37	0.27	0.66	0.22	
Number observed	3,448	3,449	3,449	3,436	3,445	3,445
Log likelihood						-1,667.98

Notes: Unrelated European ancestry subsample. Density indicates standardized natural log of population per square mile; Educated indicates proportion of those over 25 with degree; Income indicates median household income ($1,000); Home Values indicates standardized natural log of tract median home value; Miles Moved indicates the miles moved from wave 1 to wave 4; >25M indicates whether moved >25 miles from wave 1 to wave 4. All regressions account for wave 1 school and wave 4 state of residence fixed effects, and the first 20 principle components are related to ancestry; coefficients are omitted. Robust standard errors are in parentheses. [†]$p < 0.01$, *$p < 0.05$.

across the grouping variable (schools), which has analytic benefits and arguably returns more conservative estimates than pooled ordinary least squares (OLS) or fixed-effects approaches (Bell, Fairbrother, and Jones 2019). We have also presented fixed-effects estimates in our main specifications that we found to be substantively similar (see columns 5 and 6 in Tables 4 and 5). We control for cognitive performance in the first wave; sex, age, proportion of census tract that was college educated in the first wave; whether a parent attended college; and the respondent's own educational PGS. We also control for the first 20 principal components related to ancestry in our models in an effort to mitigate the likelihood that results are driven by fine-grained ancestral homophily at the local level.

In our mixed-effects logistic regression models, we find both tract-level average PGS (odds ratio [OR] = 1.15, $p = 0.006$) and spatially determined, distance-weighted

Table 2: Wave 4 neighborhood attainments; family and sibling fixed-effects models.

	Density 1k/4m^2		Proportion with college degree		Median household income	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−2.10	−2.24	0.27	0.34	84.62†	93.17†
	(1.12)	(1.26)	(0.16)	(0.18)	(27.97)	(29.25)
Education PGS	0.14†	0.14*	0.01*	0.02*	−0.14	0.84
	(0.05)	(0.06)	(0.01)	(0.01)	(1.16)	(1.44)
Degree Wave 4	0.07	0.05	0.03*	0.03	−1.12	−0.28
	(0.10)	(0.12)	(0.01)	(0.02)	(2.41)	(2.79)
Cognitive Score	0.06	−0.01	0.01	0.01	−1.49	−1.42
	(0.06)	(0.07)	(0.01)	(0.01)	(1.41)	(1.72)
Age	−0.01	−0.01	< 0.01	< 0.01	0.71	0.94
	(0.02)	(0.03)	(< 0.01)	(< 0.01)	(0.53)	(0.64)
Male	0.10	0.07	0.01	< 0.01	−0.26	0.32
	(0.07)	(0.09)	(0.01)	(0.01)	(1.83)	(2.06)
Fixed Effects	*Fam*	*Sib*	*Fam*	*Sib*	*Fam*	*Sib*
R^2	0.83	0.82	0.86	0.87	0.81	0.83
Adjusted R^2	0.47	0.45	0.55	0.60	0.41	0.47
Number observed	1,587	1,029	1,587	1,029	1,587	1,029

Notes: European ancestry subsample. Fam = Household fixed effects; Sib = Full siblings and dizygotic twins only. All regressions account for wave 4 state of residence fixed effects, coefficients are omitted. Robust standard errors are in parentheses. $^†p < 0.01$, $^*p < 0.05$.

average PGS (OR = 1.12, $p = 0.04$) to be significantly associated with college completion in wave 4 (see Table 4 and Figure 5 for a visualization of the regression lines). Thus, a single standard deviation increase in neighborhood average PGS is associated with just less than 15 percent increased odds of completing college. In our linear mixed models, however, only tract-level average PGS significantly predicted years of education ($b = 0.10$, $p = 0.014$). Because of the provocative associations between education PGS and gains in neighborhood density from adolescence to adulthood, we also stratified results on urbanicity using Add Health's indicator of whether the respondent was in a block group located in a densely settled region of 50,000 or more people as determined by the US Census Bureau (see Tables A3 and A4 in the online supplement for stratified estimates as well as Figure 5 for visualizations of these regression lines). We find that the significant associations in the main specifications for college completion are driven by respondents who lived in nonrural environments like cities and suburbs; coefficients of interest for these models were more than three times the magnitude of those based on nonurban respondents (Table A3 of the online supplement), with even greater differences among the OLS specifications (Table A4 of the online supplement).

One important consideration is whether we are proxying for friendship networks by using neighborhood and spatial renderings of average educational PGS that previous research has shown to significantly predict years of education in adulthood (Domingue et al. 2018a). To assess whether neighborhood and spatial

Table 3: Wave 4 neighborhood attainments; family and sibling fixed-effects models.

	Home Values		Miles Moved W1−W4		Moved > 25M (Logit)	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.54	−0.15	0.38	0.04	0.40	0.26
	(0.65)	(0.67)	(1.18)	(1.30)	(0.56)	(0.62)
Education PGS	0.05	0.05	0.07	0.07	−0.02	−0.02
	(0.03)	(0.03)	(0.05)	(0.06)	(0.02)	(0.03)
Degree Wave 4	0.09	0.09	0.25*	0.30*	0.21†	0.22†
	(0.06)	(0.06)	(0.10)	(0.12)	(0.05)	(0.06)
Cognitive Score	0.02	0.02	0.07	0.03	0.08†	0.06
	(0.03)	(0.04)	(0.06)	(0.08)	(0.03)	(0.04)
Age	0.02	0.02	−0.00	0.00	−0.01	−0.01
	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Male	0.07	0.06	0.10	0.06	−0.02	−0.05
	(0.04)	(0.05)	(0.08)	(0.09)	(0.04)	(0.04)
Fixed Effects	*Fam*	*Sib*	*Fam*	*Sib*	*Fam*	*Sib*
R^2	0.92	0.93	0.78	0.80		
Adjusted R^2	0.76	0.79	0.33	0.36		
Number observed	1,577	1,021	1,587	1,029	1,587	1,029
Log likelihood					271.53	178.15

Notes: European ancestry subsample. Fam = Household fixed effects; Sib = Full siblings and dizygotic twins only. All regressions account for wave (W) 4 state of residence fixed effects, coefficients are omitted. Robust standard errors are in parentheses. †$p < 0.01$, *$p < 0.05$.

measures of the genomic environment are proxying the influence of friendship networks, we devised an average PGS score based on school and home survey data for which respondents were able to elect up to 10 friends. We constructed undirected friendship networks (i.e., respondent A is considered a friend of respondent B even if their selection was not reciprocated) and took the average PGS of each individual's roster.

Running our main mixed-effects specifications with the friendship network PGS as an additional covariate, we see that they are clearly and significantly associated with educational attainment, and the coefficients for tract and buffer PGS are smaller in magnitude and not statistically significantly predictive (see Table A5 of the online supplement). When we include friend PGS in stratified models, however, we find that the significant neighborhood associations persist in the urban subgroups and that friendship circles are apparently much less influential outside of rural environments, even evincing a negative relationship with years of attainment for those raised in denser locales (Tables A6 and A7 of the online supplement). Although we cannot rule out a mechanical explanation for the divergence in the results, it seems as though friendship networks (and neighborhood conditions) are differentially influential based on the environmental context but specifically the nature of the built environment and the proximity between individuals it induces.

We also produced estimates that restricted cases to where the number of neighborhood alters used to construct the tract-level mean PGS was at least 5 or 10 (Tables
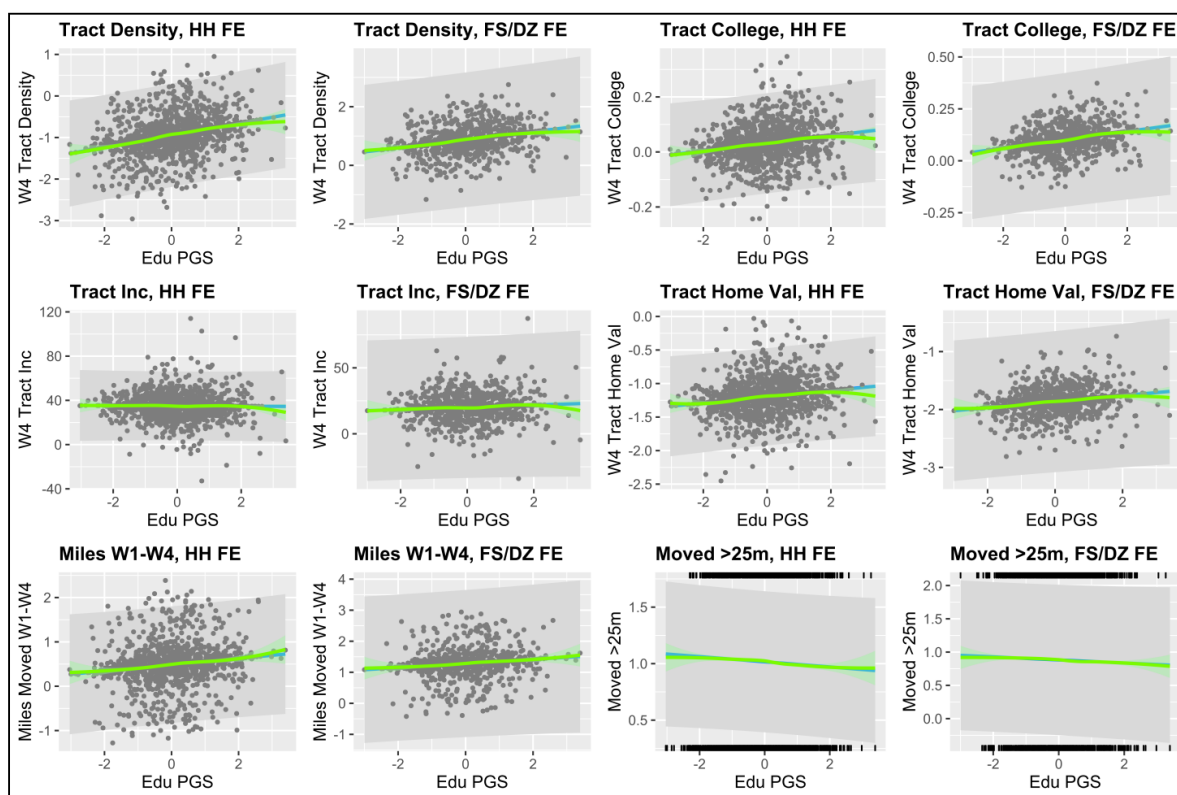
Figure 4: Visualized regression lines for individual education PGS predicting neighborhood conditions in wave 4 and distance traveled between waves (W)1 and 4 in household (HH) and full sibling/dizygotic twin (FS/DZ) fixed-effects (FE) models. Regression lines (blue for linear, green using Loess smoothing) are consistent with the estimates contained in Tables 3 and 4 and produced using the visreg package in R (see Breheny and Burchett [2017]).

A8 and A9 of the online supplement, respectively). Somewhat counterintuitively, excising cases in which relatively small numbers of neighborhood alters are used to devise mean PGS biases the sample to more rural respondents, who tend to evince larger rosters of neighborhood alters—an average of 45.5 compared to 20.26 for those in urban areas—likely because of their spatial extensiveness of census tracts in these areas. Similar to friend PGS estimates, restricting analyses to larger neighborhood alter rosters affects the main specifications, but tract-level PGS only becomes more influential in the urban subgroup. Again, tract-level PGS exhibits a negative association with years of education in the rural subgroup, though it is not significantly predictive. We also produced estimates based on neighborhood alter PGSs that were of European ancestry only and found the results were not sensitive to this qualification (see Table A10 of the online supplement).

Finally, we examined whether interactions between individual genotype and social and genetic environments were present in our data (see Tables A11 and A12 of the online supplement). We duplicated the approach for the social-genetic effect

Table 4: Degree attainment at wave 4; logit models.

	Logit		Mixed-Effects Logit		Fixed-Effects Logit	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.37	−0.10	−0.21	−0.12		
	(0.79)	(0.82)	(0.81)	(0.83)		
Cognitive Score	0.59†	0.60†	0.60†	0.61†	0.58†	0.60†
	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)
Male	−0.64†	−0.66†	−0.65†	−0.66†	−0.63†	−0.63†
	(0.08)	(0.10)	(0.08)	(0.10)	(0.09)	(0.10)
Age	−0.03	−0.04	−0.03	−0.04	−0.09*	−0.07
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Own Education PGS	0.41†	0.47†	0.41†	0.48†	0.40†	0.47†
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)
Parent College	1.31†	1.24†	1.31†	1.25†	1.23†	1.18†
	(0.10)	(0.11)	(0.10)	(0.11)	(0.09)	(0.11)
Tract College	0.31†	0.29†	0.33†	0.30†	0.34†	0.35†
	(0.05)	(0.05)	(0.05)	(0.05)	(0.09)	(0.11)
Tract Education PGS	0.15†		0.14†		0.12*	
	(0.05)		(0.05)		(0.06)	
Buffer Education PGS		0.12*		0.12*		0.11*
		(0.06)		(0.06)		(0.06)
Log likelihood	-1,610.88	-1,218.42	-1,607.17	-1,216.19	-1,344.73	-979.43
Number observed	3,194	2,409	3,194	2,409	3,194	2,409
Number of groups: school id			123	120	111	105
R^2					0.16	0.16
Maximum R^2					0.64	0.63

Notes: European ancestry subsample. All models account for the first 20 principle components related to ancestry; coefficients are omitted. Robust standard errors are in parentheses. †$p < 0.01$, *$p < 0.05$.

estimates in examining social-genetic associations using school as the grouping variable and added interaction terms between individual PGS and the proportion of the tract that was college educated along with interactions between those variables and each covariate. We found no evidence of GxE interactions in either mixed effects or naive OLS/logit specifications for either education outcome. We similarly found no evidence for the presence of SGxG, or social epistasis, when we interacted individual genome with neighborhood PGS rendered either at the tract level or as a distance-weighted measure.

Discussion

We began by examining how geographic migration, baseline neighborhood characteristics, and neighborhood attainment in adulthood related to PGS for education. First, we found suggestively large gaps in baseline neighborhood characteristics between the highest and lowest PGS quintiles and that these gaps largely grew over time as children entered adulthood. To confirm whether respondents were selecting into environments based on their genetic profile over and above other relevant characteristics (e.g., parents' education), we then used school, household,

**Table 5:** Years of education at wave 4; linear models.

	Ordinary Least Squares		Mixed Effects		Fixed Effects	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	13.36[†]	13.53[†]	13.77[†]	13.69[†]	13.95[†]	12.97[†]
	(0.58)	(0.67)	(0.65)	(0.74)	(0.90)	(1.63)
Cognitive Score	0.60[†]	0.63[†]	0.59[†]	0.63[†]	0.58[†]	0.62[†]
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
Male	−0.61[†]	−0.64[†]	−0.60[†]	−0.63[†]	−0.58[†]	−0.60[†]
	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)	(0.08)
Age	0.01	0.00	−0.01	−0.00	−0.05	−0.03
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Own Education PGS	0.35[†]	0.41[†]	0.35[†]	0.41[†]	0.34[†]	0.41[†]
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Parent College	1.34[†]	1.26[†]	1.31[†]	1.24[†]	1.25[†]	1.18[†]
	(0.09)	(0.10)	(0.09)	(0.10)	(0.09)	(0.10)
Tract College	0.28[†]	0.28[†]	0.31[†]	0.30[†]	0.35[†]	0.36[†]
	(0.04)	(0.04)	(0.05)	(0.05)	(0.07)	(0.08)
Tract Education PGS	0.11[†]		0.10[*]		0.08	
	(0.04)		(0.04)		(0.04)	
Buffer Education PGS		0.07		0.06		0.06
		(0.04)		(0.04)		(0.04)
R^2	0.29	0.29			0.34	0.34
Adjusted R^2	0.28	0.28			0.30	0.30
Number observed	3,194	2,409	3,194	2,409	3,194	2,409
Log likelihood			-6,589.99	-4,958.99		
Number of groups: school id			123	120	123	120

Notes: European ancestry subsample. All models account for the first 20 principle components related to ancestry; coefficients are omitted. Robust standard errors are in parentheses. [†]$p < 0.01$, [*]$p < 0.05$.

and sibling fixed-effects models that account for common environmental (e.g., tract characteristics) and genetic (e.g., ancestral origin) factors, and we found that PGS significantly predicted neighborhood attainments in proportion of the population with a college education and population density in adulthood across specifications.

With respect to migration, we conceived of long-distance moves as indicative of the opportunity to pursue college and labor market opportunities outside of the home metropolitan environment and thus a putatively positive outcome distinct from raw residential turnover, which could signal dislocation and tenure instability. The concrete implications of long-distance migration are somewhat difficult to ascertain, however, because of equivocal findings of the returns to interregional mobility. Current research persuasively argues for negative macroeconomic long-run effects of housing inelasticity in productive regions (Hsieh and Moretti 2019), which would imply greater gains from migrating, and highlights geographic immobility as a crucial factor in labor market disequilibrium. Even if spatial reallocation of workers results in greater aggregate levels of economic productivity, systematic migration may not necessarily deliver labor market returns on an individual level. Indeed, recent work offers no solid answers on whether interregional moves ultimately benefit movers, and the results of such work are usually sensitive to the dataset and
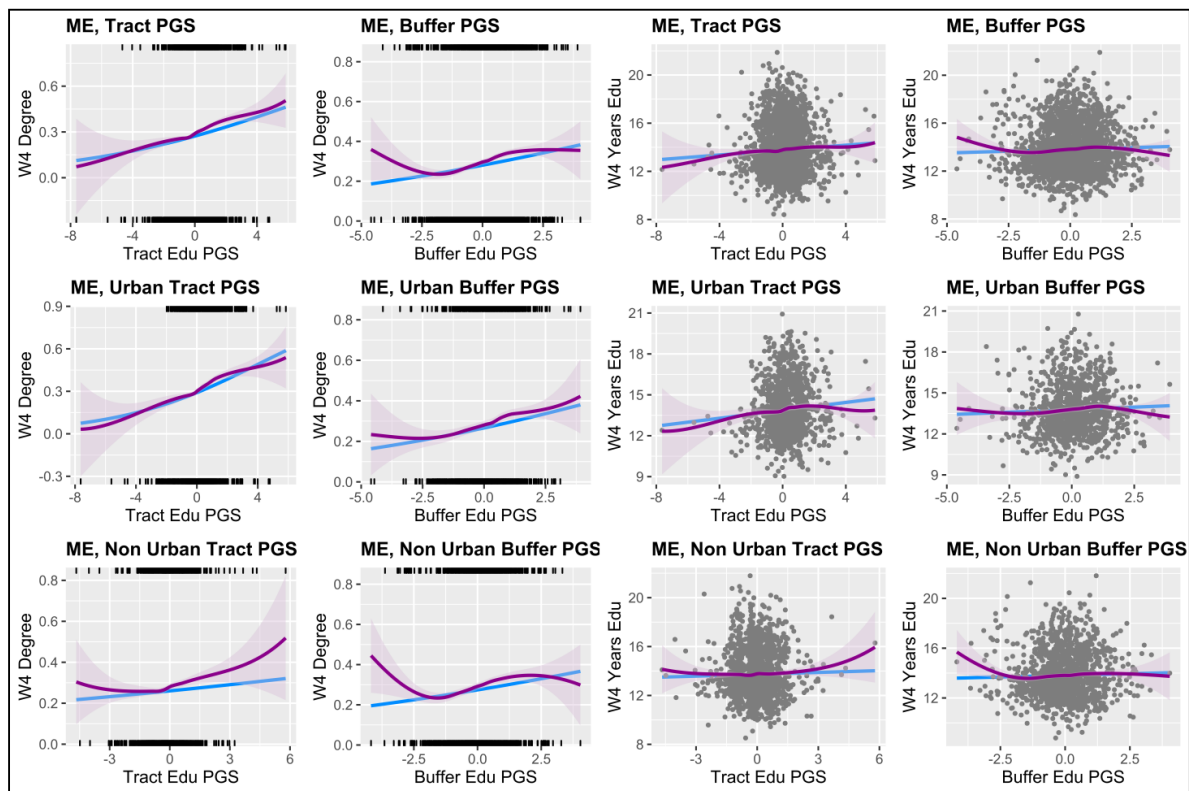
Figure 5: Visualized regression lines for mixed-effects (ME) results on tract- and buffer-based average education PGS predicting education (whether attained a college degree and years of completed education) in wave 4. Regression lines (blue for linear, purple using Loess smoothing) are consistent with the estimates contained in Table 5 and produced using the visreg package in R (see Breheny and Burchett [2017]). Edu, education; W4, wave 4.

analytic approach (Hyatt et al. 2018; Johnson and Schulhofer-Wohl 2019). Still, other research has persuasively associated regional geographic mobility with positive intergenerational mobility outcomes (Borck and Wrede 2018).

The implications of positive neighborhood attainments—for example, moving from a below- to above-average–educated census tract in the transition to adulthood—are rather more straightforward. Methodologically, active selection into different environments based on genetic background could constitute a concern for research that looks to produce causal estimates of the effects of neighborhood characteristics or for research that seeks to identify GxE interactions. In terms of real-world implications, the sorting of higher-PGS individuals into more favorable social environments could have particular significance in the presence of interaction effects or direct social-genetic effects of the chosen environment. Though we found no evidence of interaction effects in our data, we did see positive associations between neighborhood-based and spatially rendered average PGS measures and degree attainment in particular while also accounting for common school environ-

ments, the top 20 principal components related to ancestry, and a suite of other controls. We also found that in nonrural and urban environments, where associations between neighborhoods and educational outcomes is most evident, these results are robust to including measures of friends' PGS for education along with other sensitivity checks.

Although children do not actively select into neighborhood environments as they would friendship networks, there remains the possibility that parents select into neighborhood genetic environments in a way that complicates the interpretation of a putative social-genetic effect. We feel that explicitly modeling school environment as a random or fixed effect helps mitigate these concerns, as school districts would intuitively be the unit that parents care most about when selecting into new environments. We also believe that if tract-level or spatially rendered average PGS were confounded by selection and homophily, then it would be less likely that we would see differences in estimates that stratify based on the built environment. That is, if the social-genetic associations we observe were merely due to the presence of like individuals, we would be less apt to see such stark differences based on urbanicity. Intraclass correlation coefficients indicate that, if anything, there may be more genetic homophily in rural census tracts than urban ones (intraclass correlation coefficient = 0.04 vs. 0.03, respectively), though the clustering is quite modest regardless, and the overlapping standard errors render them statistically indistinguishable. It is also conceptually plausible that if a putative social-genetic effect was actually occurring and not artefactual, we would find that effect to manifest most clearly in areas where people are spatially proximate, and more opportunities for interaction or spillover might occur as a matter of course.

Still, we remain circumspect about the nature of the associations we observe and avoid speculation about mechanisms of action; future research could build on these findings by devising quasiexperimental techniques to more completely address issues of selection and confounding. We also recognize that we restricted analyses to respondents with European ancestry, which reflects a troubling lack of diversity in genetic studies (Sirugo, Williams, and Tishkoff 2019) and the weak predictive power of PGS in other subpopulations. Because we do not include many of the most economically and socially vulnerable populations in our analyses, it is unclear whether the associations we observe manifest in other subsamples or if there are relationships we are missing based on their exclusion.

Conclusion

We descriptively documented systematically different geographic mobility and neighborhood attainment based on PGS for education and found, using sibling-difference models, that genetic propensity for educational attainment suggestively predicts a subset of these outcomes over and above familial and environmental influences. We then examined the possibility for these sorting processes to magnify inequalities through social-genetic pathways and GxE and SGxG interactions. We find social-genetic associations that imply the influence of neighborhood genome but did not find evidence for interactions in our data. If neighborhood genetic environment does constitute an influential factor in outcomes of interest like educational

attainment, it could add to our understanding of spatial inequality and stratification and add weight to the implications of residential sorting that we observed.

Methodologically, active selection into different environments based on genetic background introduces a novel wrinkle in the already difficult project of accurately identifying causal peer and neighborhood effects. Indeed, questions of selection lead us to interpret our own putative social-genetic effects with caution, and future research that wields quasiexperimental analytic approaches or leverages newer data can investigate whether these associations indeed constitute a social effect. In terms of practical implications, if such social-genetic effects are not mere artifacts of selection or other methodological idiosyncrasies, then the spatial mobility and neighborhood attainment processes we document could have measurable impacts on social and economic stratification. In this sense, genetic variation at both the individual and social levels can contribute to outcomes over and above factors like parental education or wealth that have customarily been the focus of research in the social sciences. Moreover, an uneven distribution of alleles related to education could signify a new dimension of spatial inequality that is heretofore largely unexamined.

Of course, what a social (or individual) genome constitutes in the first place is a question that remains itself largely unanswered. We think it implausible that PGS for education is totally reducible to intellect or cognitive ability, and it likely captures other skills that are more nebulous but still apparently advantageous in the labor market (e.g., risk aversion and self-control [Heckman 2006]). A more critical interpretation is that PGS for education captures characteristics that are largely meaningless, like stature, physicality, or even differences in grooming, which could be transactionally valuable in certain social spheres but are mostly orthogonal to concrete skills and ability (Bourdieu 1984). Whether alleles related to higher education signify cognitive performance or merely reflect an ability to navigate the "rules of the game"—or some combination thereof—our work helps illustrate that both the individual and social genome are factors that future research should account for in mapping out spatial inequality and broader trends in social stratification that are influenced by such phenomena.

Notes

1 Sometimes alternatively referred to in the genomics literature as genetic risk scores or polygenic risk scores.

References

Abdellaoui, Abdel, David Hugh-Jones, Kathryn E. Kemper, Yan Holtz, Michel G. Nivard, Laura Veul, Loic Yengo, Brendan P. Zietsch, Timothy M. Frayling, Naomi Wray, Jian Yang, Karin J. H. Verweij, and Peter M. Visscher. 2018. "Genetic Consequences of Social Stratification in Great Britain." Retrieved October 24, 2019 https://doi.org/10.1101/457515.

Abdellaoui, Abdel, Karin J. H. Verweij, and Brendan P. Zietsch. 2014. "No Evidence for Genetic Assortative Mating Beyond That Due to Population Stratification." *Proceedings of the National Academy of Sciences* 111:E4137. https://doi.org/10.1073/pnas.1410781111.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hemous. 2019. "Innovation and Top Income Inequality." *The Review of Economic Studies* 86:1–45. https://doi.org/10.3386/w21247.

Agovino, Massimiliano, Alessandro Crociata, and Pier Luigi Sacco. 2019. "Proximity Effects in Obesity Rates in the US: A Spatial Markov Chains Approach." *Social Science and Medicine* 220:301–11. https://doi.org/10.1016/j.socscimed.2018.11.013.

Altonji, Joseph G., and Richard K. Mansfield. 2018. "Estimating Group Effects Using Averages of Observables to Control for Sorting on Unobservables: School and Neighborhood Effects." *American Economic Review* 108:2902–46. https://doi.org/10.1257/aer.20141708.

Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30:98–108. https://doi.org/10.1016/j.labeco.2014.05.008.

Arcaya, Mariana C., Reginald D. Tucker-Seeley, Rockli Kim, Alina Schnake-Mahl, Marvin So, and S. V. Subramanian. 2016. "Research on Neighborhood Effects on Health in the United States: A Systematic Review of Study Characteristics." *Social Science and Medicine* 168:16–29. https://doi.org/10.1016/j.socscimed.2016.08.047.

Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom. 2018. "Genetic Endowments and Wealth Inequality." Working Paper, National Bureau of Economic Research. https://doi.org/10.3386/w24642.

Bell, Andrew, Malcolm Fairbrother, and Kelvyn Jones. 2019. "Fixed and Random Effects Models: Making an Informed Choice." *Quality and Quantity* 53:1051–74. https://doi.org/10.1007/s11135-018-0802-x.

Belsky, Daniel W., Avshalom Caspi, Louise Arseneault, David L. Corcoran, Benjamin W. Domingue, Kathleen Mullan Harris, Renate M. Houts, Jonathan S. Mill, Terrie E. Moffitt, Joseph Prinz, Karen Sugden, Jasmin Wertz, Benjamin Williams, and Candice L. Odgers. 2019. "Genetics and the Geography of Health, Behaviour and Attainment." *Nature Human Behavior* 3:576–86. https://doi.org/10.1038/s41562-019-0562-1.

Belsky, Daniel W., Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, Avshalom Caspi, Dalton Conley, Jason M. Fletcher, Jeremy Freese, Pamela Herd, Terrie E. Moffitt, Richie Poulton, Kamil Sicinski, Jasmin Wertz, and Kathleen Mullan Harris. 2018. "Genetic Analysis of Social-Class Mobility in Five Longitudinal Studies." *Proceedings of the National Academy of Sciences* 115:E7275–84. https://doi.org/10.1073/pnas.1801238115.

Berg, Jeremy J., Arbel Harpak, Nasa Sinnott-Armstrong, Anja Moltke Joergensen, Hakhamanesh Mostafavi, Yair Field, Evan August Boyle, Xinjun Zhang, Fernando Racimo, Jonathan K. Pritchard, and Graham Coop. 2019. "Reduced Signal for Polygenic Adaptation of Height in UK Biobank." *eLife* 8:e39725. https://doi.org/10.1101/354951.

Berger, Thor, and Per Engzell. 2019. "American Geography of Opportunity Reveals European Origins." *Proceedings of the National Academy of Sciences* 116:6045–50. https://doi.org/10.1073/pnas.1810893116.

Borck, Rainald, and Matthias Wrede. 2018. "Spatial and Social Mobility." *Journal of Regional Science* 58:688–704. https://doi.org/10.1111/jors.12382.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgment of Taste.* Cambridge, MA: Harvard University Press.

Braudt, David B., and Kathleen Mullan Harris. 2018. *Polygenic Scores (PGSs) in the National Longitudinal Study of Adolescent to Adult Health (Add Health)—Release 1. Add Health Documentation*. Chapel Hill, NC: Carolina Population Center.

Brazil, Noli, and William A. V. Clark. 2017. "Residential Mobility and Dynamic Neighborhood Change During the Transition to Adulthood." *Advances in Life Course Research* 33:1–10. https://doi.org/10.1016/j.alcr.2017.06.001.

Breheny, Patrick, and Woodrow Burchett. 2017. "Visualization of Regression Lines Using Visreg." *The R Journal* 9:56–71.

Chang, Christopher C., Carson C. Chow, Laurent C. A. M. Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4:s13742-015-0047-8. https://doi.org/10.1186/s13742-015-0047-8.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *The Quarterly Journal of Economics* 133:1163–228. https://doi.org/10.3386/w23002.

Chetty, Raj, Nathaniel Hendren, Patrick M. Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129:1553–623. https://doi.org/10.1093/qje/qju022.

Conley, Dalton, Thomas Laidley, Daniel W. Belsky, Jason M. Fletcher, Jason D. Boardman, and Benjamin W. Domingue. 2016a. "Assortative Mating and Differential Fertility by Phenotype and Genotype Across the 20th Century." *Proceedings of the National Academy of Sciences* 113:6647–52. https://doi.org/10.1073/pnas.1523592113.

Conley, Dalton, Thomas Laidley, Jason D. Boardman, and Benjamin W. Domingue. 2016b. "Changing Polygenic Penetrance on Phenotypes in the 20th Century Among Adults in the US Population." *Scientific Reports* 6:30348. https://doi.org/10.1038/srep30348.

Domingue, Benjamin W., Daniel W. Belsky, Jason M. Fletcher, Dalton Conley, Jason D. Boardman, and Kathleen Mullan Harris. 2018a. "The Social Genome of Friends and Schoolmates in the National Longitudinal Study of Adolescent to Adult Health." *Proceedings of the National Academy of Sciences* 115:702–7. https://doi.org/10.1101/107045.

Domingue, Benjamin W., David H. Rehkopf, Dalton Conley, and Jason D. Boardman. 2018b. "Geographic Clustering of Polygenic Scores at Different Stages of the Life Course." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 4:137–49. https://doi.org/10.7758/rsf.2018.4.4.08.

Domingue, Benjamin W., Jason Fletcher, Dalton Conley, and Jason D. Boardman. 2014. "Genetic and Educational Assortative Mating Among US Adults." *Proceedings of the National Academy of Sciences* 111:7996–8000. https://doi.org/10.1073/pnas.1321426111.

Feld, Jan, and Ulf Zölitz. 2017. "Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects." *Journal of Labor Economics* 35:387–428. https://doi.org/10.1086/689472.

Galster, George, Anna Santiago, Lisa Stack, and Jackie Cutsinger. 2016. "Neighborhood Effects on Secondary School Performance of Latino and African American Youth: Evidence from a Natural Experiment in Denver." *Journal of Urban Economics* 93:30–48. https://doi.org/10.1016/j.jue.2016.02.004.

Graham, Bryan S. 2018. "Identifying and Estimating Neighborhood Effects." *Journal of Economic Literature* 56:450–500.

Hamer, Darryl, and Leah Sirota. 2000. "Beware the Chopsticks Gene." *Molecular Psychiatry* 5:11–13. https://doi.org/10.1038/sj.mp.4000662.

Haworth, Simon, Ruth Mitchell, Laura Corbin, Kaitlin H. Wade, Tom Dudding, Ashley Budu-Aggrey, David Carslake, Gibran Hemani, Lavinia Paternoster, George Davey Smith, Neil Davies, Daniel J. Lawson, and Nicholas J. Timpson. 2019. "Apparent Latent Structure within the UK Biobank Sample has Implications for Epidemiological Analysis." *Nature Communications* 10:333. https://doi.org/10.1038/s41467-018-08219-1.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312:1900–02. https://doi.org/10.1126/science.1128898.

Hicks, Andrew L., Mark S. Handcock, Narayan Sastry, and Anne R. Pebley. 2018. "Sequential Neighborhood Effects: The Effect of Long-Term Exposure to Concentrated Disadvantage on Children's Reading and Math Test Scores." *Demography* 55:1–31. https://doi.org/10.1007/s13524-017-0636-5.

Highland, Heather M., Christy L. Avery, Qing Duan, Yun Li, and Kathleen Mullan Harris. 2018. *Quality Control Analysis of Add Health GWAS Data. Add Health Documentation.* Chapel Hill, NC: Carolina Population Center.

Hout, Michael. 2012. "Social and Economic Returns to College Education in the United States." *Annual Review of Sociology* 38:379–400. https://doi.org/10.1146/annurev.soc.012809.102503.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11:1–39. https://doi.org/10.1257/mac.20170388.

Huang, Ying, Scott J. South, and Amy Spring. 2017. "Racial Differences in Neighborhood Attainment: The Contributions of Interneighborhood Migration and In Situ Change." *Demography* 54:1819–43. https://doi.org/10.1007/s13524-017-0606-y.

Hugh-Jones, David, Karin J. H. Verweij, Beate St. Pourcain, and Abdel Abdellaoui. 2016. "Assortative Mating on Educational Attainment Leads to Genetic Spousal Resemblance for Polygenic Scores." *Intelligence* 59:103–8. https://doi.org/10.1016/j.intell.2016.08.005.

Hyatt, Henry, Erika McEntarfer, Ken Ueda, and Alexandria Zhang. 2018. "Interstate Migration and Employer-to-Employer Transitions in the United States: New Evidence from Administrative Records Data." *Demography* 55:2161–80. https://doi.org/10.1007/s13524-018-0720-5.

Johnson, Janna E., and Sam Schulhofer-Wohl. 2019. "Changing Patterns of Geographic Mobility and the Labor Market for Young Adults." *Journal of Labor Economics* 37:S199–241. https://doi.org/10.1086/700887.

Keller, Matthew C. 2014. "Gene × Environment Interaction Studies have not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution." *Biological Psychiatry* 75:18–24. https://doi.org/10.1016/j.biopsych.2013.09.006.

Kerminen, Sini, Aki S. Havulinna, Garrett Hellenthal, Alicia R. Martin, Antti-Pekka Sarin, Markus Perola, Aarno Palotie, Veikko Salomaa, Mark J. Daly, Samuli Ripatti, and Matti Pirinen. 2017. "Fine-Scale Genetic Structure in Finland." *G3: Genes, Genomes, Genetics* 7:3459–68. https://doi.org/10.1534/g3.117.300217.

Kleinepier, Tom, and Maarten van Ham. 2018. "The Temporal Dynamics of Neighborhood Disadvantage in Childhood and Subsequent Problem Behavior in Adolescence." *Journal of Youth and Adolescence* 47:1611–28. https://doi.org/10.1007/s10964-018-0878-6.

Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni J. Vilhjalmsson, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, Asmundur Oddsson, Bjarni V. Halldorsson, Gisli Masson, Daniel F. Gudbjartsson, Agnar Helgason, Gyda Bjornsdottir,

Unnur Thorsteindottir, and Kari Stefansson. 2018. "The Nature of Nurture: Effects of Parental Genotypes." *Science* 359:424–8. https://doi.org/10.1126/science.aan6877.

Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, Peter Bowers, Julia Sidorenko, Richard Karlsson Linnér, Mark Alan Fontana, Tushar Kundu, Chanwook Lee, Hui Li, Ruoxi Li, Rebecca Royer, Pascal N. Timshel, Raymond K. Walters, Emily A. Willoughby, Loïc Yengo, 23andMe Research Team, COGENT (Cognitive Genomics Consortium), Social Science Genetic Association Consortium, Maris Alver, Yanchun Bao, David W. Clark, Felix R. Day, Nicholas A. Furlotte, Peter K. Joshi, Kathryn E. Kemper, Aaron Kleinman, Claudia Langenberg, Reedik Mägi, Joey W. Trampush, Shefali Setia Verma, Yang Wu, Max Lam, Jing Hua Zhao, Zhili Zheng, Jason D. Boardman, Harry Campbell, Jeremy Freese, Kathleen Mullan Harris, Caroline Hayward, Pamela Herd, Meena Kumari, Todd Lencz, Jian'an Luan, Anil K. Malhotra, Andres Metspalu, Lili Milani, Ken K. Ong, John R. B. Perry, David J. Porteous, Marylyn D. Ritchie, Melissa C. Smart, Blair H. Smith, Joyce Y. Tung, Nicholas J. Wareham, James F. Wilson, Jonathan P. Beauchamp, Dalton C. Conley, Tõnu Esko, Steven F. Lehrer, Patrik K. E. Magnusson, Sven Oskarsson, Tune H. Pers, Matthew R. Robinson, Kevin Thom, Chelsea Watson, Christopher F. Chabris, Michelle N. Meyer, David I. Laibson, Jian Yang, Magnus Johannesson, Philipp D. Koellinger, Patrick Turley, Peter M. Visscher, Daniel J. Benjamin, and David Cesarini. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50:1112–21. https://doi.org/10.1038/s41588-018-0147-3.

Leibbrand, Christine, Catherine Massey, J. Trent Alexander, and Stewart Tolnay. 2019. "Neighborhood Attainment Outcomes for Children of the Great Migration." *American Journal of Sociology* 125:141–83. https://doi.org/10.1086/703682.

Levy, Brian L. 2019. "Heterogeneous Impacts of Concentrated Poverty During Adolescence on College Outcomes." *Social Forces* 98:147–82. https://doi.org/10.1093/sf/soy116.

Levy, Brian L., Ann Owens, and Robert J. Sampson. 2019. "The Varying Effects of Neighborhood Disadvantage on College Graduation: Moderating and Mediating Mechanisms." *Sociology of Education* 92:269–92. https://doi.org/10.1177/0038040719850146.

Liu, Hexuan. 2018. "Social and Genetic Pathways in Multigenerational Transmission of Educational Attainment." *American Journal of Sociology* 83:278–304. https://doi.org/10.1177/0003122418759651.

Minh, Anita, Nazeem Muhajarine, Magdalena Janus, Marni Brownell, and Martin Guhn. 2017. "A Review of Neighborhood Effects and Early Child Development: How, Where, and for Whom, Do Neighborhoods Matter?" *Health and Place* 46:155–74. https://doi.org/10.1016/j.healthplace.2017.04.012.

Novembre, John, Toby Johnson, Katarzyna Bryc, Zoltán Kutalik, Adam R. Boyko, Adam Auton, Amit Indap, Karen S. King, Sven Bergmann, Matthew R. Nelson, Matthew Stephens, and Carlos D. Bustamante. 2008. "Genes Mirror Geography within Europe." *Nature* 456:98–101. https://doi.org/10.1038/nature07331.

Pais, Jeremy. 2017. "Intergenerational Neighborhood Attainment and the Legacy of Racial Residential Segregation: A Causal Mediation Analysis." *Demography* 54:1221–50. https://doi.org/10.1007/s13524-017-0597-8.

Rehkopf, David H., Benjamin W. Domingue, and Mark R. Cullen. 2016. "The Geographic Distribution of Genetic Risk as Compared to Social Risk for Chronic Diseases in the United States." *Biodemography and Social Biology* 62:126–42. https://doi.org/10.1080/19485565.2016.1141353.

Sampson, Robert J. 2019. "Neighbourhood Effects and Beyond: Explaining the Paradoxes of Inequality in the Changing American Metropolis." *Urban Studies* 56:3–32. https://doi.org/10.1177/0042098018795363.

Sampson, Robert J., and Patrick Sharkey. 2008. "Neighborhood Selection and the Social Reproduction of Concentrated Racial Inequality." *Demography* 45:1–29. https://doi.org/10.1353/dem.2008.0012.

Sazzini, Marco, Guido Alberto Gnecchi Ruscone, Cristina Giuliani, Stefania Sarno, Andrea Quagliariello, Sara De Fanti, Alessio Boattini, Davide Gentilini, Giovanni Fiorito, Mariagrazia Catanoso, Luigi Boiardi, Stefania Croci, Pierluigi Macchioni, Vilma Mantovani, Anna Maria Di Blasio, Giuseppe Matullo, Carlo Salvarani, Claudio Franceschi, Davide Pettener, Paolo Garagnani, and Donata Luiselli. 2016. "Complex Interplay between Neutral and Adaptive Evolution Shaped Differential Genomic Background and Disease Susceptibility along the Italian Peninsula." *Scientific Reports* 6:32513. https://doi.org/10.1038/srep32513.

Sirugo, Giorgio, Scott M. Williams, and Sarah A. Tishkoff. 2019. "The Missing Diversity in Human Genetic Studies." *Cell* 177:26–31. https://doi.org/10.1016/j.cell.2019.02.048.

Sotoudeh, Ramina, Dalton Conley, and Kathleen Mullan Harris. 2017. "The Influence of Peer Genotypes and Behavior on Smoking Outcomes: Evidence from Add Health." Working Paper, National Bureau of Economic Research. https://doi.org/10.3386/w24113.

South, Scott J., Ying Huang, Amy Spring, and Kyle Crowder. 2016. "Neighborhood Attainment over the Adult Life Course." *American Sociological Review* 81:1276–304. https://doi.org/10.1177/0003122416673029.

Torche, Florencia. 2011. "Is a College Degree Still the Great Equalizer? Intergenerational Mobility Across Levels of Schooling in the United States." *American Journal of Sociology* 117:763–807. https://doi.org/10.1086/661904.

Tran, Van C. 2019. "Second-Generation Contextual Mobility: Neighborhood Attainment from Birth to Young Adulthood in the United States." *International Migration Review,* first published on April 3, 2019 as https://doi.org/10.1177/0197918319832235.

Trejo, Sam, Daniel W. Belsky, Jason D. Boardman, Jeremy Freese, Kathleen Mullan Harris, Pam Herd, Kamil Sicinski, and Benjamin W. Domingue. 2018. "Schools as Moderators of Genetic Associations with Life Course Attainments: Evidence from the WLS and Add Health." *Sociological Science* 5:513–40. https://doi.org/10.15195/v5.a22.

van Ham, Maarten, Sanne Boschman, and Matt Vogel. 2018. "Incorporating Neighborhood Choice in a Model of Neighborhood Effects on Income." *Demography* 55:1069–90. https://doi.org/10.1007/s13524-018-0672-9.

Wang, Chaolong, Sebastian Zöllner, and Noah A. Rosenberg. 2012. "A Quantitative Comparison of the Similarity between Genes and Geography in Worldwide Human Populations." *PLOS Genetics* 8:e1002886. https://doi.org/10.1371/journal.pgen.1002886.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass and Public Policy*. Chicago, IL: University of Chicago Press.

Wodtke, Geoffrey T. 2013. "Duration and Timing of Exposure to Neighborhood Poverty and the Risk of Adolescent Parenthood." *Demography* 50:1765–88. https://doi.org/10.1007/s13524-013-0219-z.

Wodtke, Geoffrey T., Felix Elwert, and David J. Harding. 2016. "Neighborhood Effect Heterogeneity by Family Income and Developmental Period." *American Journal of Sociology* 121:1168–222. https://doi.org/10.1086/684137.

Yang, Tse-Chuan, and Scott J. South. 2018. "Neighborhood Effects on Body Mass: Temporal and Spatial Dimensions." *Social Science and Medicine* 217:45–54. https://doi.org/10.1016/j.socscimed.2018.09.055.

Thomas Laidley: Institute of Behavioral Science, University of Colorado Boulder.
E-mail: thla0691@colorado.edu.

Justin Vinneau: Institute of Behavioral Science, University of Colorado Boulder.
E-mail: justin.vinneau@colorado.edu.

Jason D. Boardman: Institute of Behavioral Science, University of Colorado Boulder.
E-mail: boardman@colorado.edu.