Journal of Health and Social Behavior

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT

to article in

Journal of Health and Social Behavior, 2021, Vol. 62, Issue 2

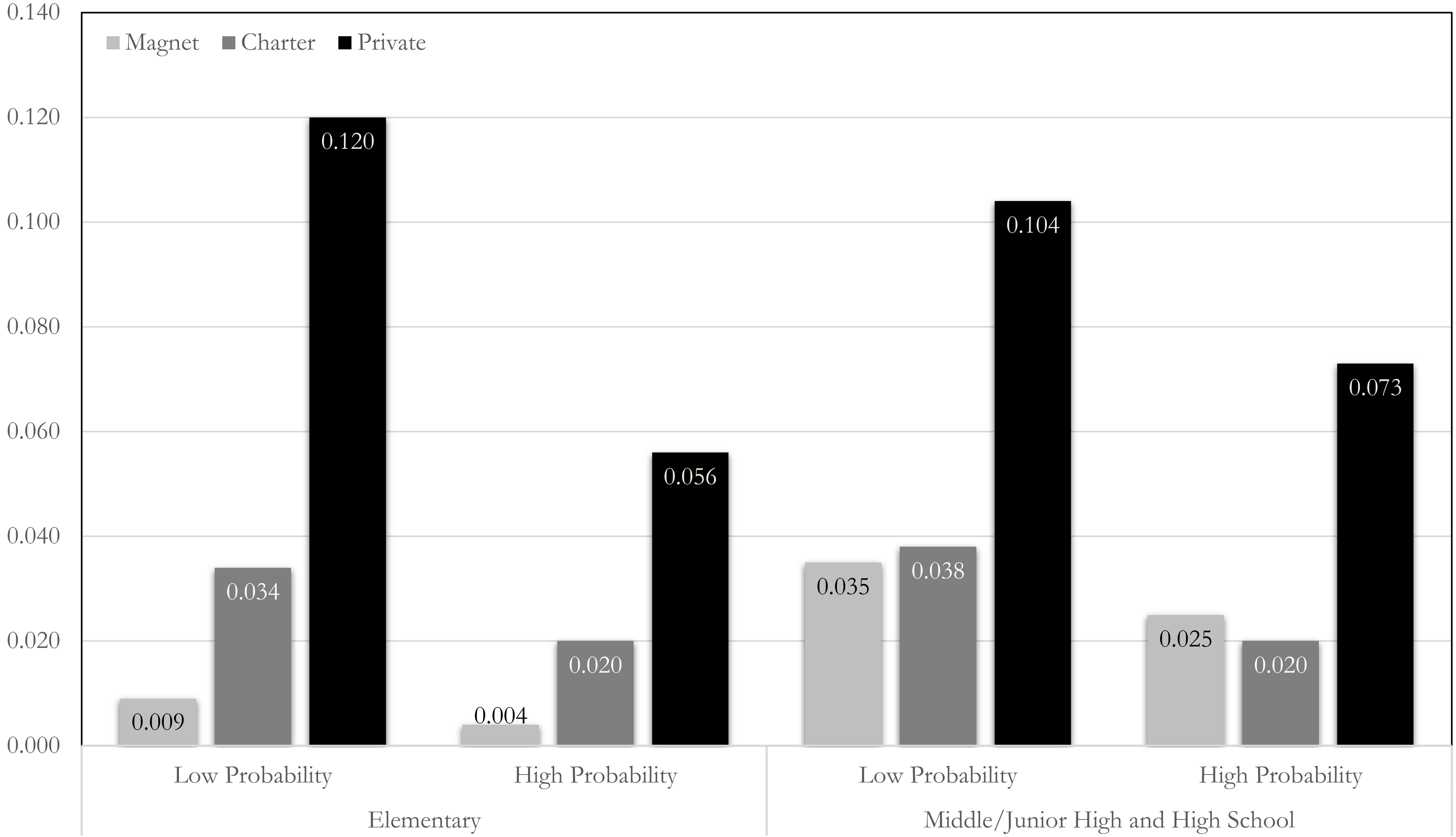
Parental Depression and Contextual Selection: The Case of School Choice

Jared N. Schachner
University of Chicago

APPENDIX A. SUPPLEMENTARY TABLES AND FIGURES

FIGURE A1

Unadjusted Probability of L.A.FANS Child Enrolling in a Magnet, Charter, or Private School by School Level and Parental Depression



Note:

[a] Estimates are based on L.A.FANS Pooled Child-Wave Sample (Wave 1: 2000–2002, Wave 2: 2006–2008), with analytic weights.

[b] Low depression probability is defined as .5 or lower; high probability is defined as over .5.

TABLE A1

Effects of L.A.FANS Child, Parent, and Household Characteristics on Magnet, Charter, or Private School Enrollment, Logit Models

Analytic Sample	Model 1: All Children		Model 2: All Children		Model 3: All Children		Model 4: Latino/Black Only		Model 5: White/Asian/Other	
Variables	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
PCG likelihood of depression			–.943**	.326	–.936**	.345	–1.008*	.478	–.623	.494
Child race										
Asian	–.208	.508	–.263	.491	.163	.419			–.141	.495
Latino	–.274	.342	–.324	.341	–.036	.532				
Black	–.311	.450	–.332	.444	–.169	.610	–.343	.698		
Other/Multiracial	–.425	.478	–.444	.473	–.039	.672			–.193	.602
Parent/household attributes										
PCG first generation immigrant	.085	.203	.096	.197	–.401	.245	–.691+	.383	–.234	.412
Household income (log)	.406*	.160	.387*	.162	.458*	.195	.434	.281	.534*	.236
Homeowner	.019	.289	.023	.289	.331	.366	–.235	.583	1.106*	.512
PCG completed some college	.824**	.201	.826**	.201	1.026**	.266	1.243**	.365	.937+	.493
PCG Bachelor's degree+	1.407**	.284	1.389**	.287	1.124**	.283	1.901**	.557	.755	.564
PCG marital status: married	–.134	.215	–.133	.213	–.164	.278	.210	.368	–.383	.676
Number of children in household	.038	.090	.030	.090	–.044	.101	.026	.111	–.370	.245
Constant	–1.742**	.556	–1.595**	.555	–.652	.666	–2.366*	1.183	1.407	1.133
Neighborhood Fixed Effects	N		N		Y		Y		Y	
PCG Depression AME	N/A		–.109		–.095		–.097		–.077	
Household *N*	1,678		1,678		1,437		878		445	
Child *N*	1,678		1,678		1,437		878		445	

Note:

[a] PCG = L.A.FANS primary caregiver.

[b] All models contain analytic weights accounting for L.A.FANS sampling/attrition, a control for child age, and fixed effects capturing survey wave (2006–2008, ref: 2000–2002), child gender, and school level.

[c] Standard errors are clustered by Los Angeles County neighborhood of residence.

[d] + $p < .10$, *$p < .05$, **$p < .01$ (two–tailed test).

TABLE A2

Heterogeneous Effects of L.A.FANS Child, Parent, and Household Characteristics on Magnet, Charter, or Private School Enrollment by Age, Logit Models

Analytic Sample	Model 1 All Races: Elementary		Model 2: All Races: Junior High/High		Model 3 All Races: All Ages		Model 4 Latino/Black: Elementary		Model 5 Latino/Black: Junior High/High		Model 6: Latino/Black: All Ages	
Variables	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
PCG likelihood of depression	–1.092+	.658	–.568	.434	–.456	.360	–1.788*	.780	–.960*	.424	–.618	.387
PCG depression X Elementary					–.430	.501					–.820	.633
Child attributes												
Elementary school age					–.056	.375					.070	.459
Asian	.569	.703	–.004	.477	.204	.444						
Latino	–.877	.630	–.274	.544	–.433	.465						
Black	–.056	.659	–.213	.588	–.091	.498	.831	.945	–.026	.685	.130	.541
Other/Multiracial	–.419	.863	–.671	.458	–.488	.501						
Latino/Black												
Parent/household attributes												
PCG first generation immigrant	–.486	.394	.030	.316	–.211	.269	–.664	.825	.303	.437	–.096	.459
Household income (log)	.362	.280	.392*	.188	.380*	.184	.507	.394	.346	.273	.379+	.223
Homeowner	.981+	.533	.447	.400	.637+	.330	.946	.830	–.088	.585	.320	.519
PCG completed some college	1.143**	.386	.629+	.324	.815**	.250	1.337*	.590	.751+	.441	1.137**	.349
PCG Bachelor's degree+	1.360*	.566	.888*	.440	.996**	.282	1.906**	.701	1.309+	.685	1.649**	.449
PCG marital status: married	–.397	.405	.042	.334	–.183	.260	–.025	.559	–.196	.351	–.169	.322
Number of children in household	.052	.154	–.010	.081	.024	.087	.105	.218	–.030	.089	.084	.116
Constant	–.857	.986	–.752	1.007	–.372	.858	–3.655*	1.613	–1.224	1.697	–2.077+	1.182
PCG Depression AME$_{Elementary}$	–.107				–.087		–.142				–.131	
PCG Depression AME$_{Junior\ High/High}$			–.064		–.047				–.097		–.056	
Child-Year *N*	1,003		1,237		2,489		588		749		1,568	

Note:

[a] PCG = L.A.FANS primary caregiver.

[b] All models contain analytic weights accounting for L.A.FANS sampling/attrition, a control for child age, and fixed effects capturing survey wave (2006–2008, ref: 2000–2002), child gender, and neighborhood.

[c] Standard errors are clustered by Los Angeles County neighborhood of residence.

[d] + $p < .10$, *$p < .05$, **$p < .01$ (two–tailed test).

APPENDIX B. METHODOLOGICAL DETAILS

Sample Specification and Missing Data

As the main text notes, 2,906 L.A.FANS child-wave combinations qualify for the analytic sample because the RSC or SIB were ages 5–17, enrolled in neither college nor special education, and a complete child survey, valid school enrollment information, census tract geocodes, and analytic weights accounting for sampling design and attrition (the latter for wave 2 observations only) were available.

Of these, 37 child-waves (1%) were missing a wave-specific PCG depression probability estimate, leaving 2,869 child-wave combinations that contain the core dependent and independent variable. Of these 2,869 observations, 115 (4%) were missing values on one or more control variables. The only variables for which there are missing values are: household income (logged) (2%), and indicators for homeownership (0.02%), as well as PCG's educational attainment (0.05%), marital status (0.05%), and nativity status (0.07%).

Given that these missing data rates are very low and appear to be missing at random, for my core analyses I conducted listwise deletion of the 115 cases with missing values on control variables and ran all core models on the complete data covering 2,754 child-wave combinations, 2,247 unique child respondents, and 1,683 unique primary caregivers/households. However, replication analyses that include the 115 child-wave observations with one or more missing values, and apply imputed values for each, are available upon request.

Assigning Child-Wave Observations to Magnet Schools versus Traditional Public Schools

Identifying which L.A.FANS analytic sample children attended either a magnet or traditional public school in a given wave is not as straightforward as assigning children to private or charter schools. Many Los Angeles County magnet schools share campuses with traditional public schools and therefore do not receive a unique school identification code from the state. However, California's school directory does indicate whether a given school campus contains a co-resident magnet school. Thus, I could safely assume all children attending a non-charter public school without a co-resident magnet were traditional public school attendees, and I marked them as "0."

The final subset of children attended a public school containing a co-resident magnet program which they may or may not have attended. Within this group, if the PCG reported her child attended a magnet program during the wave in question, I marked the child as “1”, indicating she was a magnet student and therefore enrolled in a school of choice. All remaining children were marked as “0” because there was no evidence they attended a private, charter, or magnet school.

Operationalizing Depression

My core independent variable – probability of depression – is drawn from the Composite International Diagnostic Interview-Short Form (CIDI-SF), which was developed by the World Health Organization, based on the Diagnostic and Statistical Manual of Mental Disorders, and used in the U.S. National Health Interview Survey, as well as the Fragile Families Survey. The CIDI-SF is constructed to generate two separate scores pertaining to major depressive episodes: one capturing dysphoric symptoms and the other anehedonic symptoms.

All respondents are asked if they felt sad, blue, or depressed for two weeks or more in the past year. If they respond yes, then a set questions intended to gauge dysphoric symptoms are asked (e.g., how long during the day sadness lasts, how often in two weeks the respondent felt sad, whether she lost interest in things, felt more tired than usual, gained/lost weight, had trouble concentrating or sleeping). If they respond no, they skip the remaining dysphoric symptom questions and are asked a question related to anehedonic symptoms.

This question asks respondents if they have lost interest in hobbies, work, or activities that usually provide pleasure. If they respond yes, then a set of questions intended to gauge anehedonic symptoms is asked (e.g., whether or not she felt worthless, thought about death). Note that all respondents who report experiencing dysphoric symptoms also proceed to answer if they have lost interest in activities, and if they respond yes to this too, then they also answer questions related to anehedonic symptoms. If a respondent reports that she is not experiencing dysphoric symptoms nor has she lost interest in activities, she has completed the CIDI-SF and is reported to have no dysphoric or anehedonic symptoms.

The dual structure of the questionnaire generates two scores: one quantifying the number of dysphoric symptoms (0–7) and the other quantifying the number of anehedonic symptoms (0–6), based on the answers to each question. L.A.FANS translates these symptom scores into three probability measures: one capturing the probability of a major depressive episode marked by dysphoria, the second capturing the probability of a major depressive episode marked by anehedonia, and the third capturing the probability of a major depressive episode of either type. This latter probability is the one I use as my core predictor in all models.

Operationalizing Spatial Fixed Effects

Social scientists typically operationalize neighborhoods as census tracts, which are U.S. Census-defined geographic areas containing 1,200 to 8,000 residents that are redrawn every decade. Because L.A.FANS generated its sample, in part, based on household residence within 65 randomly-selected census tracts in Los Angeles County and because the survey tracks each respondent's tract at wave 1 and 2, it seems logical to employ this spatial unit when attempting to capture neighborhood fixed effects in this study.

However, the census tract has important conceptual limitations in general, and empirical limitations for this study, in particular. Although census tracts are small spatial units that can capture fine-grained variation in social conditions between different clusters of city blocks, the census tract itself is not symbolically meaningful as a spatial unit. The vast majority of people do not know what their residential census tract boundaries are, people do not select housing units based on census tract location, and people's physical activity spaces typically extend far beyond their residential census tract on a daily basis. Urban scholars have long sought symbolically salient neighborhood units for American metropolitan areas, but there is no standardized method for constructing them, especially because metros vary widely in their spatial structures. The census' standardized spatial definition has thus become the default for social scientists, despite these conceptual limitations.

However, in some metropolitan areas like Chicago and Los Angeles County, researchers and organizations have taken it upon themselves to develop distinct sets of neighborhood boundaries that fit their unique ecological context and convey symbolic meaning to local residents. Around this study's timeframe, 2009–2010, the *Los Angeles Times*' Mapping L.A. project did just this. A group of *Times* reporters and web developers in 2009–10 developed a set of neighborhood boundaries and used crowdsourced input to refine them. The current Mapping L.A. boundaries are available online in ArcGIS format for anyone to download for free (see https://maps.latimes.com/about/). The boundaries are broadly perceived as properly capturing most Angelenos' perceptions, and the neighborhoods contain names and reputations that will be familiar to locals and non-locals alike (e.g., Westwood, Beverly Hills). Many local government agencies and nonprofit organizations have adopted these boundaries to report spatial variation in social conditions across the vast county.

The empirical benefits of using Mapping L.A. neighborhood-based fixed effects, rather than tract fixed effects, for this particular study are even clearer and more compelling than the conceptual benefits of doing so. Because there are over 2,000 census tracts in Los Angeles County, but only 272 Mapping L.A. neighborhoods, I can cluster my analytic sample of child-waves into a smaller set of spatial units by using the former: 130 communities rather than ~340 census tracts. As a result, the average cluster size is much larger (~21 versus ~8), which ensures fewer cases are dropped from models due to a lack of intra-cluster variation in the school sorting outcome. Moreover, this larger spatial fixed effect yields more stable results, especially when examining racial heterogeneity in depression's effects.

As mentioned above, L.A.FANS data link all child-waves to census tract locations. However, they do not assign child-waves to Mapping L.A. neighborhood locations. In order to assign each child-wave to an appropriate Mapping L.A. neighborhood, so that I could construct spatial fixed effects based on them, I used ArcGIS software to execute a spatial overlay of the *Los Angeles Times*-provided Mapping L.A. boundaries with census tracts, in 2000 boundaries. The vast majority of tracts are fully subsumed within Mapping L.A. neighborhoods. However, for the tracts that span multiple neighborhoods, I assigned the tract to the neighborhood that the majority of its land area covers. Thus, for each child-wave in my dataset, the neighborhood fixed effect captures the Mapping L.A. neighborhood that the majority of its census tract's land area covers. More details on this procedure, and the rationale underlying it, are available upon request.

Using AMEs to Gauge Interaction Effects in Logistic Regressions

Unlike traditional linear models, regression models predicting non-continuous outcomes do not generate easily-interpretable coefficients on interaction terms. The intuition is that when applying model specifications relying on Maximum Likelihood Estimates that generate predicted probabilities (e.g., logistic regression models), the size and significance of group-based differences in the outcome "depend on the values of the regressors where the comparison is made" (Long and Mustillo 2018). Group-based comparisons of odds ratios can also prove misleading. Thus, an emerging consensus—underscored in a recent article in *American Sociological Review*—advises against using interaction terms' coefficients and standard errors generated from logistic regression models to draw any conclusions (Mustillo, Lizardo, and McVeigh 2018). Thus, the size and significance of interaction effects must be estimated through alternative strategies.

Mize (2019) and Long and Mustillo (2018) converge on one promising strategy: comparing group differences in regressors' marginal effects on the predicted probability of the outcome. Three types of marginal effects are typically estimated: Discrete Change at Representative (DCR) Values; Marginal Effects at the Means (MEM; sometimes referred to Discrete Change at the Mean, DCM); and Average Marginal Effects (AME; sometimes referred to as Average Discrete Change).

MEM refers to a marginal effect of a given predictor for the "average person" in a given sample (Mize 2019). Concretely, MEM estimates compare the change in predicted probabilities of the outcome if the predictor of interest increases by a discrete change (e.g., one standard deviation) for a hypothetical observation in which all other variables are held at the sample mean.

Mize (2019) points out that this hypothetical observation may not be a realistic representation of the actual observations in the sample (e.g., no observation may have mean levels of all variables). AMEs, on the other hand, sidestep this issue by estimating marginal effects based on actual, rather than hypothetical, observations. AMEs are estimated by calculating the marginal effect of increasing a given variable by a discrete amount for each individual observation on its predicted probability of the outcome, while keeping all other variables' observed values unchanged, and then averaging these marginal effects across the entire analytic sample.

To gauge interaction effects, AMEs can be calculated separately by subsample and then their magnitudes can be compared. This is exactly the approach I pursued in my tests of Hypothesis #2 presented in Table 3: I estimated the AMEs of parental depression on the predicted probability of school choice activation for Whites, Latinos, and Blacks separately and then employed a Wald test to assess whether these group differences in depression's AME are significant.

Using California's Similar Schools Ranking to Gauge School Quality

As mentioned in the main text, I used California's Similar Schools Ranking of public schools as a proxy for the quality of each school attended by children in my analytic sample. This ranking is drawn from the California Department of Education's Academic Performance Index (API) reporting system, which tracks demographic and test score data for every public school campus with eleven or more valid scores, every year between 1998 and 2013. Within this database, a schoolwide API score, ranging from a low of 200 to a high of 1000, was generated based on all students' *levels* of performance on standardized tests (aggregated across reading, math, and other subjects). The California Department of Education reported this score as an "absolute" measure and also generated two types of statewide "relative" rankings based on it: (1) API Statewide Rank and (2) API Similar Schools Ranking.

The API Statewide Rank merely ranked all schools of the same level (e.g., elementary, middle/junior, high) in the entire state and assigned each school a decile (1–10) based on this ranking, with 10 indicating the school scored in the top 10% of all state schools. Instead of ranking all schools in the entire state relative to each other, the API Similar Schools Ranking attempted to only rank the API scores of schools relative to other schools with *similar socio-demographic characteristics*. Although the methodology for calculating the peer group against which each school would be ranked for its 1–10 Similar Schools Ranking is complex, the main intuition is that this ranking operated as a kind of value-added measure that attempted to isolate school performance from the influence of race and class composition differences across campuses that could explain why some schools performed better than others. See the 2011-12 Academic Performance Index Reports Information Guide (https://edsource.org/wp-content/uploads/old/API-explanation20121.pdf) for more details on the methodology underlying the Similar Schools Ranking. Also note that the Similar Schools Ranking is publicly disclosed via the Internet and newspapers, rendering it accessible to parents and the public. For savvy parents seeking to maximize academic quality it is this particular value-added type ranking that should in theory drive school enrollment decisions.