# Supporting Information

# Changes in exposure to ambient fine particulate matter due to relocation and long-term survival in Canada: a quasi-experimental study

Hong Chen,[1,2,3,4] Jay S. Kaufman,[5,6] Toyib Olaniyan,[7] Lauren Pinault,[7] Michael Tjepkema,[7] Li Chen,[1] Aaron van Donkelaar,[8] Randall V. Martin,[9] Perry Hystad,[10] Chen Chen,[12] Megan Kirby-McGregor,[5] Li Bai,[3] Richard T. Burnett,[1] Tarik Benmarhnia[11,12]

[1] Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, Canada
[2] Public Health Ontario, Toronto, ON, Canada
[3] ICES, Toronto, ON, Canada
[4] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[5] Department of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada
[6] Institute for Health and Social Policy, McGill University, Montreal, QC, Canada
[7] Health Analysis Division, Statistics Canada, Ottawa, ON, Canada
[8] Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada
[9] Department of Energy, Environment & Chemical Engineering, Washington University, St Louis, Missouri, United States
[10] College of Public Health and Human Sciences, Oregon State University, Corvallis, USA
[11] Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA
[12] Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, CA, USA

**Correspondence:**

Hong Chen, PhD
Environmental Health Science and Research Bureau
Health Canada
101 Tunney's Pasture
Ottawa, Ontario K1A 0K9
Tel: 343-542-3483
Email: hong.chen@hc-sc.gc.ca

# Table of Contents

***Study Population***

In this quasi-experimental study, we identified movers from participants in the CanCHEC, a large national cohort in Canada. Details of CanCHEC have been presented elsewhere.[1] Briefly, CanCHEC comprises respondents to the Canadian long-form census questionnaires that collect data on demographic and socioeconomic characteristics from one-in-five randomly selected Canadian households from 1991 until 2006. Many of these characteristics are determinants of residential mobility.[1]

Using standard deterministic and probabilistic record linkage techniques with information on names (if available), birth date, sex, residential postal code, and marital status, Statistics Canada linked census respondents to both Canadian family income tax files, which provide annual information on all households' economic characteristics and residential postal codes,[2] and Canadian Institute of Health Information's hospital abstract discharge database, which provides hospital records from acute-care hospitals in all Canadian provinces and territories except for the province of Quebec.[3] Using the social insurance number, a unique identifier for all Canadians, Statistics Canada also linked the respondents to the national vital statistics death database that contains all reported deaths in Canada including death date and underlying cause of death. The linkage rates vary from 78.6% to 99.8% (depending on the database).[2,4] CanCHEC has been frequently used to examine the $PM_{2.5}$-related health effects.[2,4-6]

***Outcomes***

The primary outcome of interest was nonaccidental death (International Classification of Diseases, Ninth Revision ICD-9 code: <800 and ICD-10 code: A00-R99). We also considered three secondary mortality outcomes including deaths from any cardiometabolic cause (ICD-9:

390-459, 250 and ICD-10: I00-I99, E10-E14),[7] any respiratory cause (ICD-9: 460-519 and ICD-10: J00–J99), and all cancer causes (ICD-9: 140-239; ICD-10: C00–D48). All outcomes were obtained from the national vital statistics database.

***Covariates***

We obtained the following data from cohort members' responses to the census questionnaire: age; sex; race/ethnicity (*i.e.*, visible minority status and Indigenous identity); nativity; marital status; educational attainment; occupational class; and employment status. Education attainment was defined as less than high school, high school, post-secondary non-university, or university. Occupational class was categorized as management, professional, skilled, semi-skilled, unskilled, or not applicable (designating those not in the labor force). Using family income tax files, we also derived annual household income adequacy (in deciles) which accounted for household income, family size, region, and year. In addition, to characterize cohort members' baseline health status, we derived the Charlson comorbidity score, an index commonly used to measure prognosis for mortality and disease burden, using hospitalization data over three years preceding the baseline.[8] These variables influence mortality and are known to affect residential mobility.[9]

To further account for regional differences in mortality that may not be attributable to air pollution, we derived neighborhood-level deprivation based on the Canadian Marginalization Index which was previously developed to characterize inequalities in health and social wellbeing.[10] Like previous studies,[4,5] we defined four deprivation variables, one for each dimension underlying the construct of marginalization (residential instability, material deprivation, dependency, and ethnic concentration), based on census tracts in cities and census subdivisions (*i.e.*, municipalities) outside of larger metropolitan areas. Additionally, we created

an urban form variable to characterize active commuting and transit-use using census tract data, given reports of active commuting improving fitness and health.[11] In addition, considering growing evidence linking residential green spaces to mortality,[12,13] we obtained the satellite-derived Normalized Difference Vegetation Index (NDVI), an objective measure of vegetative greenness.[12,13] As done previously,[14,15] we calculated NDVI annual measures for a 500-m buffer area around each residential postal code. Furthermore, to control for regional differences in access to health services, we measured distance to the nearest healthcare facility (including family-physician offices, hospitals, and community health centres) from each postal code and we created a dichotomous variable indicating whether the move led to a change in the administrative health region.[16] All the area-level variables were defined at baseline.

***Ascertainment of Exposure to $NO_2$ and $O_3$***

To explore whether exposure to other air pollutants, especially nitrogen dioxide ($NO_2$) and ozone ($O_3$) might explain the $PM_{2.5}$-mortality association, we estimated ambient concentrations of $NO_2$ and $O_3$ using a national land-use regression (LUR) model and an optimal interpolation technique, respectively. The national LUR of $NO_2$ was developed from fixed-site monitoring data, satellite $NO_2$ estimates, area of industrial land use, road length, and mean summer rainfall.[17] At a spatial resolution of 100 $m^2$, the LUR model explained 73% of the variation in annual 2006 measurements of $NO_2$, with a RMSE (root mean square error) of 2.9 parts per billion (ppb).[17] To estimate $O_3$ exposure, Environment and Climate Change Canada applied the optimal interpolation technique that combines the true observations of $O_3$ with chemical transport models that account for meteorological and chemical patterns of $O_3$, thus producing a long-term annual mean warm-season exposure surface of $O_3$ (21 ×21 km) covering Canada between 2002 and

2009.[18] For both pollutants, we derived annual exposures at each residential postal code at baseline, after applying similar temporal adjustments.[17,18]

***Propensity Score Matching and Analysis***

To emulate a hypothetical randomized experiment in which eligible subjects can be randomly assigned to either a high or a low (or intermediate) $PM_{2.5}$ exposure group, followed by comparing their mean mortality rates over five years, we conducted a propensity score matching analysis.[19] The propensity score for the probability of moving from a high to a low or intermediate $PM_{2.5}$ postal code was estimated for each individual in the high $PM_{2.5}$ cohort using a logistic regression model with all demographic, socioeconomic, health, and environmentally related covariates including attained age, sex, race/ethnicity, indigenous identity, nativity, marital status, education, occupation, employment status, household income adequacy, Charlson comorbidity index, residential proximity to healthcare services, an indicator for changing health region or not, residential greenness, $NO_2$ and $O_3$, urban form characteristics, and neighborhood-level dependency, material deprivation, residential instability, and ethnic concentration. In addition, we included participant's prior exposure to $PM_{2.5}$ over five years preceding baseline, airshed at baseline (lived in East Central airshed or not), and the index year of moving due to a concern that the likelihood of moving and mortality risk may vary over time.[2,4] These variables were selected *a priori* for inclusion because they may potentially confound the association between $PM_{2.5}$ and mortality according to the literature.[5,8,9] Analogously, we estimated a propensity score for the probability of moving from a low to a high (or intermediate) $PM_{2.5}$ postal code for each individual in the low $PM_{2.5}$ cohort.

Specifically, we estimated propensity score using equation (1) as follows:

$$\text{logit } pr(\text{high-to-low vs. high-to-high } PM_{2.5}) = \boldsymbol{V} + \boldsymbol{L} \quad (1)$$

where $\boldsymbol{V}$ is a vector of personal-level variables measured on census reference day including sex, race/ethnicity, indigenous identity, immigrant status, marital status, educational attainment, occupational class, and employment status. $\boldsymbol{L}$ is a vector of variables measured in the year of moving (baseline) including attended age (in five-year age groups), annual family income adequacy (in decile), Charlson comorbidity index over 3 years prior, long-term exposure to $PM_{2.5}$ over five years prior, indicators for calendar year of moving, indicator for a change in the administrative health region (yes/no), and according to the destination area, airshed, proximity to healthcare services (in quintile), urban form characteristics, annual mean $NO_2$, annual mean $O_3$, neighborhood-level dependency (in quintile), deprivation (in quintile), instability (in quintile), and ethnic composition (in quintile), as well as residential greenness (NDVI). Analogously, we constructed propensity score models for all other changes in $PM_{2.5}$ of interest.

We matched each individual who moved to a different $PM_{2.5}$ exposure group (the exposed) to up to three individuals who moved within the same exposure group (the control). For example, for each high-to-low mover, we matched up to three high-to-high movers. A nearest-neighbor matching without replacement was applied to match individuals on the basis of the logit of their propensity score, with a caliper of 0.2.[20,21] We assessed the balance in the distribution of covariates before and after matching using standardized differences, with a difference of <0.1 after matching considered a good balance.[22] The propensity-score estimation and matching were done for each cohort separately.

To assess the relationship between changes in $PM_{2.5}$ and mortality, we used Cox proportional hazards models with matching weights applied and time-on-study (in days) as the time scale. Like previous studies using propensity score matching,[23-25] exposure was represented by an indicator variable for exposure groups (*e.g.*, high-to-low versus high-to-high). As a secondary measure of changes in $PM_{2.5}$, we fitted an interaction between the group indicator and the difference in annual mean concentrations of $PM_{2.5}$ between an individual's origin and destination. In all Cox models, we adjusted for all covariates used for propensity score matching, because double adjustment allowed for removing confounding that may arise from any residual imbalance after matching.[26] To further account for the paired nature of the matched cohort, we used robust sandwich-type variance estimators to construct valid 95% confidence intervals (CIs) of the estimated hazard ratio, which was expressed as the mean percent change in mortality (Δ%).[27]

Specifically, to assess the relationship between changes in $PM_{2.5}$ and mortality, we fitted Cox models using equation (2) as follows:

$$h(t) = h_0(t)\exp(\beta \times \boldsymbol{A} + \gamma \times \boldsymbol{V} + \psi \times \boldsymbol{L}) \quad (2)$$

where **A** is an indicator for a change in $PM_{2.5}$ groups (*e.g.*, 1=high-to-low and 0=high-to-high), and ***V*** *and* ***L*** refer to the same set of variables shown in Equation #1 with corresponding coefficients of $\beta$, $\gamma$, and $\psi$.

In addition, as a secondary measure of changes in $PM_{2.5}$, we fitted alternative Cox models using equation (3) as follows:

$$h(t) = h_0(t)\exp(\beta \times \boldsymbol{A} \times \Delta PM_{2.5} + \gamma \times \boldsymbol{V} + \psi \times \boldsymbol{L}) \quad (3)$$

where $\Delta PM_{2.5}$ denotes the difference in annual mean concentrations of $PM_{2.5}$ between an individual's origin and destination and all other notations remain the same as in Equation #2.

To maintain covariate balance in the matched cohorts, we conducted our primary analysis according to the exposure group that an individual was initially "assigned" at baseline, regardless of any departure from that exposure during the 5-year follow-up. This is a close analogy to the intention-to-treat principle used extensively in randomized experiments which provides unbiased effect estimates when there is noncompliance with initial assignment.[28,29] As a supportive analysis, we also conducted the per-protocol analysis in which we censored individuals at the time when they moved to a different exposure group during follow-up (~12%).

***Assessment of the Implication of Missing Data Using Multiple Imputation***

To further examine the implications of missing on our results, we have carried out a series of new analyses. Multiple imputation is a widely used technique to handle complex incomplete data problems in health research.[30-37] It has been shown in numerous studies to produce valid inferences and it allows for incorporating statistical uncertainty about the value to impute.[38,39] Therefore, we applied multiple imputation in an effort to examine a range of possible inferences given missing data for $NO_2$, income decile, and some other covariates. In doing this, we explicitly imposed missing at random (MAR) as the missing mechanism for all covariates.

Previous studies[40,41] have shown that using all observed information in multiple imputation reduced bias and maximized certainty. Failure to do so may yield bias in survival analysis after imputation.[42] As a result, to impute missing data, we used all available information.

Our procedures for creating multiple imputation and conducting post-imputation analysis consisted of three steps. In the first step, we specified an imputation model separately for each covariate with missing values: (1) for continuous variables such as $NO_2$ and $O_3$, we used predictive mean matching, a semi-parametric imputation method that allows for the preservation of non-linear relations.[43] In a recent simulation study comparing the performance of different imputation methods for handling missing covariate data in survival analysis,[44] predictive mean matching was found to produce less biased estimates than linear models in imputing continuous variables; and (2) for ordered categorical variables such as material deprivation (lowest, lower-middle, middle, upper-middle, and upper income quintile), income decile, and Charlson comorbidity score (0, 1, 2, 3 or higher), we used a proportional odds model.[45]

Due to computational constraints arising from the large size of our analytical cohort, we were able to implement only a limited number of iterations for multiple imputation. Previous simulation work using moderate amounts of missing data yielded satisfactory performance with three to five iterations.[46] As a result, we repeated the imputation to obtain five copies of the "filled-in" dataset. In the subsequent survival analyses, the effect estimates using these five imputed datasets were broadly consistent, indicating that five iterations were sufficient.

To include the outcome as a predictor in multiple imputation, we applied the approach by White and Royston (2009).[47] In a simulation study, White and Royston demonstrated that comparing to an earlier approach considering survival time after a logarithmic transformation (log(T)), using cumulative baseline hazard $H_0(T)$ yielded minimal bias in imputation with survival data.[47] Following White and Royston (2009),[47] we thus derived $H_0(T)$ using the Nelson-Aalen estimator and included $H_0(T)$ in the imputation models.

In step 2, using each imputed copy of the dataset, we first conducted propensity score matching and then fitted Cox proportional hazards model, adjusting for the same covariates as described in our main text. The analysis was repeated for each of the five imputed copies.

In step 3, we combined the five analyses by computing the mean of hazard ratios and a standard error that accounts for the average variability observed both *within* and *between* these separate analyses, using a pooling method by Rubin (1987).[39]

We implemented multiple imputation using the MICE (multiple imputation by chained equations) library of R statistical package (version 3.6.2).[48]

**SI Appendix Table S2** shows the % change in mortality estimated from each of the five filled-in datasets, a pooled estimate, and the estimate from our original analysis based on complete data. The effect estimates were virtually unchanged between these analyses, suggesting that our risk estimates were not appreciably influenced by missing information on material deprivation, $NO_2$, and all other covariates.

This finding is consistent with a recent simulation study which showed that with up to 10% of multivariate missingness with a MAR mechanism, survival analysis using complete data provided reasonable estimates of hazard ratios and associated standard errors.[44]

***Additional sensitivity analyses***

We have conducted three additional sensitivity analyses. In the first analysis, we conducted a separate analysis for each of the two age groups (younger age: 25-64 years and older age: >65 years). In the second analysis, due to the concern over the possibility of moving to retirement

homes, we created a new variable indicating whether the destination area has a retirement home. Because information on retirement residence was unavailable at the individual-level, we created an area-level variable based on the publicly available postal-code information of retirement homes across Canada (*e.g.*, Ontario's Retirement Home Database), and we then adjusted for this new variable in the survival analysis. Furthermore, we considered age (in months) as an alternative time scale in Cox models and we repeated the analysis.

As shown in Table S5, the associations between changes in $PM_{2.5}$ and mortality were broadly similar between the two age groups. In the second analysis where we further adjusted for the indicator for moving to a destination area with a retirement home, the $PM_{2.5}$-related effects on mortality remained consistent. In the third analysis in which we used age as the time scale, the effect estimates were largely unchanged.

***Syntax of main statistical codes***

*Propensity score models*

```
m.out <- matchit(treatment ~ factor(age_bin) + factor(sex) + factory(year) + PM25_5yr +
factor(Charlson) + …, data=indata, method="nearest", ratio=3, distance="logit", caliper=0.2,
replace=F)
```

*Cox models*

cox.out <- coxph(surv(survival, nonaccdeath) ~ factor(treatment) + factor(age_bin) + factor(sex) + factory(year) + PM25_5yr + factor(Charlson) + …, weights=weights, robust=T, cluster=subclass, data=match.data(m.out))  # ITT for high-to-low vs. high-to-high PM2.5

cox.out <- coxph(surv(survival, nonaccdeath) ~ factor(treatment):I(L_PM25-B_PM25) + factor(age_bin) + factor(sex) + factory(year) + …, weights=weights, robust=T, cluster=subclass, data=match.data(m.out))  # ITT for high-to-low vs. high-to-high PM2.5 (expressed as per μg/$m^3$)

cox.out <- coxph(surv(survival_censored_postmove, nonaccdeath_censored_postmove) ~ factor(treatment):I(L_PM25-B_PM25) + factor(age_bin) + factor(sex) + factory(year) + …, weights=weights, robust=T, cluster=subclass, data=match.data(m.out)) ## Per protocol for high-to-low vs. high-to-high PM2.5

*Multiple imputation*

ch <- nelsonaalen(indata, survival, nonaccdeath)

indata$ch <- ch

indata.imputed <- mice(indata, m=5, seed=1)

**Table S1.** Baseline characteristics of matched cohorts (mean ± standard deviation or percent, otherwise specified) *

| Baseline Characteristics | Matched high-to-moderate group † | Matched high-to-low group † | Matched low-to-moderate group † | Matched low-to-high group † |
|---|---|---|---|---|
| | (N=157,985) | (N=25,310) | (N=112,650) | (N=15,940) |
| ***Demographic characteristics*** | | | | |
| Age, y | 46.8±15.6 | 48.7±16.1 | 49.8±16.2 | 51.0±16.3 |
| Men | 52 | 52 | 52 | 52 |
| Race/ethnicity | | | | |
| White or Indigenous | 73 | 78 | 96 | 95 |
| Visible minority | 27 | 22 | 4 | 5 |
| Indigenous identity | | | | |
| Not Indigenous | 99 | 99 | 93 | 95 |
| Aboriginal | 1 | 1 | 7 | 5 |
| Landed immigrant | 44 | 38 | 12 | 16 |
| Marital status | | | | |
| Single | 17 | 16 | 12 | 13 |
| Common-law | 6 | 7 | 8 | 7 |
| Married | 62 | 62 | 66 | 65 |
| Separated | 3 | 3 | 3 | 3 |
| Divorced | 6 | 6 | 5 | 5 |
| Widowed | 6 | 7 | 6 | 7 |
| Education | | | | |
| Less than high school | 22 | 21 | 29 | 26 |
| High school | 33 | 33 | 37 | 35 |
| Post-secondary non-university | 22 | 23 | 21 | 22 |
| University | 23 | 23 | 13 | 17 |
| Employment | | | | |
| Employed | 72 | 70 | 63 | 62 |
| Unemployed | 4 | 3 | 5 | 6 |
| Not in labor force | 24 | 27 | 32 | 32 |
| Occupation | | | | |
| Management | 9 | 10 | 8 | 8 |
| Professional | 14 | 15 | 11 | 12 |
| Skilled | 21 | 22 | 24 | 23 |
| Semi-skilled | 27 | 25 | 23 | 22 |
| Unskilled | 7 | 6 | 8 | 7 |
| Not applicable | 22 | 22 | 27 | 28 |
| Household income adequacy | | | | |
| 10th decile - lowest | 11 | 14 | 11 | 11 |

| Baseline Characteristics | Matched high-to-moderate group † | Matched high-to-low group † | Matched low-to-moderate group † | Matched low-to-high group † |
|---|---|---|---|---|
| | (N=157,985) | (N=25,310) | (N=112,650) | (N=15,940) |
| 9th decile | 11 | 12 | 11 | 10 |
| 8th decile | 11 | 12 | 10 | 11 |
| 7th decile | 11 | 10 | 10 | 9 |
| 6th decile | 11 | 11 | 10 | 10 |
| 5th decile | 10 | 9 | 10 | 10 |
| 4th decile | 10 | 9 | 10 | 11 |
| 3rd decile | 9 | 9 | 10 | 10 |
| 2nd decile | 9 | 10 | 9 | 10 |
| 1st decile - highest | 7 | 10 | 8 | 8 |
| ***Clinical characteristics*** | | | | |
| Charlson comorbidity index over three years prior | | | | |
| 0 | 94 | 93 | 92 | 91 |
| 1 | 2 | 3 | 4 | 4 |
| 2 | 2 | 2 | 2 | 2 |
| 3 or higher | 2 | 2 | 2 | 3 |
| ***Environmental characteristics*** | | | | |
| Annual $PM_{2.5}$ averaged over five years prior ($\mu g/m^3$) | 11.7±2.0 | 11.1±1.9 | 4.7±0.8 | 4.6±0.8 |
| Baseline annual mean $NO_2$ (ppb) | 12.3±3.7 | 9.4±3.0 | 7.1±3.7 | 9.1±4.1 |
| Baseline annual mean $O_3$ (ppb) | 45.7±6.3 | 44.8±6.9 | 34.1±7.1 | 38.0±7.3 |
| Baseline greenness within a 500-m buffer around home | 0.5±0.1 | 0.5±0.1 | 0.5±0.1 | 0.5±0.1 |
| Urban form characteristics | | | | |
| Active urban core | 4 | 2 | 2 | 6 |
| Transit-reliant suburb | 4 | 1 | 1 | 4 |
| Car-reliant suburb | 68 | 54 | 30 | 45 |
| Exurban | 7 | 14 | 4 | 3 |
| Non-CMA/CA ‡ | 16 | 30 | 63 | 42 |
| Lived in East Central Airshed | 96 | 91 | 20 | 38 |
| ***Healthcare and social-economic characteristics*** | | | | |
| Moved to another health region | 35 | 51 | 32 | 50 |
| Proximity to healthcare services | | | | |
| ≤ 3 km | | | | |
| 1st quintile - lowest | 22 | 31 | 7 | 5 |
| 2nd quintile | 21 | 21 | 13 | 8 |
| 3rd quintile | 20 | 15 | 19 | 16 |
| 4th quintile | 14 | 10 | 19 | 23 |
| 5th quintile - highest | 8 | 4 | 13 | 25 |

| Baseline Characteristics | Matched high-to-moderate group † | Matched high-to-low group † | Matched low-to-moderate group † | Matched low-to-high group † |
|---|---|---|---|---|
| | (N=157,985) | (N=25,310) | (N=112,650) | (N=15,940) |
| > 3 km | 15 | 19 | 29 | 22 |
| Dependency | | | | |
| 1st quintile - lowest | 32 | 26 | 20 | 21 |
| 2nd quintile | 25 | 23 | 15 | 15 |
| 3rd quintile | 17 | 18 | 14 | 16 |
| 4th quintile | 14 | 17 | 21 | 21 |
| 5th quintile - highest | 12 | 16 | 30 | 27 |
| Material deprivation | | | | |
| 1st quintile - lowest | 36 | 47 | 22 | 23 |
| 2nd quintile | 25 | 25 | 24 | 22 |
| 3rd quintile | 19 | 17 | 21 | 22 |
| 4th quintile | 10 | 9 | 17 | 20 |
| 5th quintile - highest | 9 | 3 | 16 | 14 |
| Residential instability | | | | |
| 1st quintile - lowest | 45 | 44 | 16 | 16 |
| 2nd quintile | 28 | 34 | 26 | 25 |
| 3rd quintile | 12 | 14 | 27 | 23 |
| 4th quintile | 8 | 5 | 22 | 22 |
| 5th quintile - highest | 6 | 3 | 9 | 14 |
| Ethnic concentration | | | | |
| 1st quintile - lowest | 17 | 22 | 28 | 23 |
| 2nd quintile | 20 | 25 | 29 | 28 |
| 3rd quintile | 16 | 16 | 22 | 23 |
| 4th quintile | 18 | 15 | 12 | 16 |
| 5th quintile - highest | 30 | 23 | 8 | 11 |

* All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

† Defined by tertiles of annual mean $PM_{2.5}$ in the cohort.

‡ CMA/CA: census metropolitan area/census agglomeration area.

**Table S2.** Sensitivity analysis of associations between changes in $PM_{2.5}$ exposure due to residential mobility (defined by tertiles) and the risk of nonaccidental mortality in Canada, 1997 to 2016, using multiple imputation (N=5)

| | **Moving from high to intermediate $PM_{2.5}$** * | | **Moving from high to low $PM_{2.5}$** * | | **Moving from low to intermediate $PM_{2.5}$** † | | **Moving from low to high $PM_{2.5}$** † | |
|---|---|---|---|---|---|---|---|---|
| **Sensitivity analysis** | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) |
| ***Multiple imputation*** | | | | | | | | |
| 1 | 7,815 | -6.5% (-11.3% to -1.5%) | 1,645 | -10.8% (-20.8% to 0.3%) | 7,505 | 1.6% (-3.6% to 7.1%) | 1,195 | 13.9% (-0.2% to 30.0%) |
| 2 | 7,855 | -7.6% (-12.4% to -2.6%) | 1,675 | -13.3% (-22.9% to -2.5%) | 7,495 | 1.9% (-3.4% to 7.4%) | 1,200 | 11.2% (-2.7% to 27.1%) |
| 3 | 7,830 | -6.0% (-10.8% to -1.0%) | 1,640 | -10.1% (-20.1% to 1.2%) | 7,455 | 1.9% (-3.4% to 7.4%) | 1,180 | 15.6% (1.1% to 32.2%) |
| 4 | 7,850 | -4.4% (-9.3% to 0.8%) | 1,635 | -11.8% (-21.7% to -0.6%) | 7,510 | 3.0% (-2.3% to 8.6%) | 1,210 | 10.2% (-3.5% to 26.0%) |
| 5 | 7,835 | -8.7% (-13.3% to -3.7%) | 1,635 | -14.6% (-24.2% to -3.7%) | 7,510 | 2.8% (-2.5% to 8.4%) | 1,215 | 14.9% (0.7% to 31.1%) |
| ***Pooled estimate*** ‡ | 7,840 | -6.7% (-12.5% to -0.4%) | 1,645 | -12.1% (-22.6% to -0.2%) | 7,495 | 2.2% (-3.2% to 8.0%) | 1,200 | 13.2% (-1.6% to 30.2%) |
| ***Complete-data analysis*** § | 7,435 | -6.8% (-11.7% to -1.7%) | 1,535 | -12.8% (-23.0% to -1.3%) | 6,960 | 1.2% (-4.1% to 6.9%) | 1,075 | 13.2% (-1.5% to 30.2%) |

* Reference level is moving from high to high $PM_{2.5}$ environments (defined by the upper tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

† Reference level is moving from low to low $PM_{2.5}$ environments (defined by the lower tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

‡ Summary estimate was obtained using Rubin's rule. The number of events was averaged the five iterations. Due to computational constraints, we were only able to conduct multiple imputation with five realizations (each realization required up to two days).

§ The same results in Table 2 are presented here, to ease comparison.

**Table S3.** Sensitivity analysis of associations between changes in $PM_{2.5}$ exposure due to residential mobility (defined by tertiles) and the risk of nonaccidental mortality in Canada, 1997 to 2016, using the unmatched cohorts and standard Cox regression models

| | | | Comparing exposure groups | | |
|---|---|---|---|---|---|
| **Exposure contrast** | **Number of subjects** | **Number of events** | % change in mortality | Lower confidence limit | Upper confidence limit |
| High to intermediate * | | | | | |
| + age, sex, and year | 385,380 | 19,785 | -3.4% | -7.1% | 0.4% |
| + personal level variables† | 385,380 | 19,785 | -5.3% | -8.9% | -1.6% |
| + environmental variables‡ | 385,380 | 19,785 | -8.1% | -11.9% | -4.2% |
| + other area-level variables | 385,380 | 19,785 | -11.5% | -15.3% | -7.5% |
| High to low * | | | | | |
| + age, sex, and year | 333,925 | 18,035 | -2.4% | -7.6% | 3.2% |
| + personal level variables† | 333,925 | 18,035 | -5.7% | -10.8% | -0.3% |
| + environmental variables‡ | 333,925 | 18,035 | -8.4% | -14.8% | -1.5% |
| + other area-level variables§ | 333,925 | 18,035 | -14.3% | -20.3% | -7.9% |
| Low to intermediate * | | | | | |
| + age, sex, and year | 245,375 | 14,025 | -5.1% | -8.6% | -1.4% |
| + personal level variables† | 245,375 | 14,025 | -0.5% | -4.3% | 3.3% |
| + environmental variables‡ | 245,375 | 14,025 | -0.6% | -5.2% | 4.3% |
| + other area-level variables§ | 245,375 | 14,025 | 0.2% | -4.5% | 5.1% |
| Low to high * | | | | | |
| + age, sex, and year | 200,750 | 11,275 | 6.5% | -0.6% | 14.2% |
| + personal level variables† | 200,750 | 11,275 | 8.3% | 0.9% | 16.1% |
| + environmental variables‡ | 200,750 | 11,275 | 8.4% | -1.2% | 19.0% |
| + other area-level variables§ | 200,750 | 11,275 | 9.1% | -0.9% | 20.0% |

* Reference level for high to low or intermediate $PM_{2.5}$ environments is moving from high to high $PM_{2.5}$ environments (defined by the upper tertile of annual mean $PM_{2.5}$). Reference level for moving from low to high or intermediate $PM_{2.5}$ environments is moving from low to low $PM_{2.5}$ environments (defined by the lower tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

† Additionally adjusted for race/ethnicity, indigenous identity, immigrant status, marital status, educational attainment, occupation, employment status, annual family income, and Charlson comorbidity index.

‡ Additionally adjusted for environmental variables including residential greenness, $NO_2$, $O_3$, and urban form characteristics.

§ Additionally adjusted for all other area-level variables including residential proximity to healthcare services, an indicator for changing health region or not, airshed, and four neighborhood-level variables about dependency, material deprivation, residential instability, and ethnic concentration.

**Table S4.** Sensitivity analysis of associations between changes in $PM_{2.5}$ exposure due to residential mobility (defined by tertiles) and the risk of nonaccidental mortality in Canada, 1997 to 2016, using the study cohorts including movers who participated multiple censuses and propensity score matching analysis

| | **Number of** | **Number of** | Comparing exposure groups | | |
|---|---|---|---|---|---|
| **Exposure contrast** | **subjects** | **events** | % change in mortality | Lower confidence limit | Upper confidence limit |
| High to intermediate * | 179,945 | 9,140 | -6.5% | -10.9% | -1.8% |
| High to low * | 29,735 | 1,825 | -10.1% | -19.7% | 0.6% |
| Low to intermediate † | 138,070 | 9,145 | 3.8% | -1.1% | 8.9% |
| Low to high † | 19,875 | 1,425 | 21.2% | 7.4% | 36.8% |

* Reference level is moving from high to high $PM_{2.5}$ environments (defined by the upper tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

† Reference level is moving from low to low $PM_{2.5}$ environments (defined by the lower tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

**Table S5.** Additional sensitivity analyses of associations between changes in $PM_{2.5}$ exposure due to residential mobility (defined by tertiles) and the risk of nonaccidental mortality in Canada, 1997 to 2016

| | **Moving from high to intermediate $PM_{2.5}$** * | | **Moving from high to low $PM_{2.5}$** * | | **Moving from low to intermediate $PM_{2.5}$** † | | **Moving from low to high $PM_{2.5}$** † | |
|---|---|---|---|---|---|---|---|---|
| **Sensitivity analysis** | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) | # of events | Comparing exposure groups (% change, 95% CI) |
| ***Age at baseline*** | | | | | | | | |
| 25-64 years | 1,210 | -9.0% (-20.4% to 3.9%) | 245 | -12.4% (-36.5% to 21.0%) | 1,270 | 4.4% (-8% to 18.5%) | 185 | 9.8% (-22.4% to 55.2%) |
| ≥ 65 years | 6,225 | -6.3% (-11.7% to -0.6%) | 1,290 | -12.8% (-24.0% to 0.1%) | 5,690 | 0.7% (-5.1% to 6.9%) | 890 | 13.9% (-2.4% to 32.9%) |
| *Interaction p-value* | | 0.69 | | 0.98 | | 0.62 | | 0.85 |
| ***Further adjusted for whether moving to an area with a retirement home*** ‡ | 7,435 | -7.2% (-12.1% to -2.1%) | 1,535 | -12.8% (-23.1% to -1.2%) | 6,960 | 1.0% (-4.3% to 6.7%) | 1,075 | 12.2% (-2.5% to 29.1%) |
| ***Used age as the time scale*** § | 7,435 | -6.3% (-11.5% to -0.9%) | 1,535 | -14.3% (-25.0% to -2.0%) | 6,960 | 0.9% (-4.8% to 6.8%) | 1,075 | 10.6% (-2.8% to 28.5%) |
| ***Original analysis*** ‖ | 7,435 | -6.8% (-11.7% to -1.7%) | 1,535 | -12.8% (-23.0% to -1.3%) | 6,960 | 1.2% (-4.1% to 6.9%) | 1,075 | 13.2% (-1.5% to 30.2%) |

* Reference level is moving from high to high $PM_{2.5}$ environments (defined by the upper tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

† Reference level is moving from low to low $PM_{2.5}$ environments (defined by the lower tertile of annual mean $PM_{2.5}$). All counts were rounded up to the nearest five in compliance with privacy requirements by Statistics Canada.

‡ 1: moved to a postal-code destination area with a retirement home; 0: otherwise

§ Age (in months) was used as the time scale in Cox models.

‖ The same results in Table 2 of main text are presented here, to ease comparison.

**Figure Legends**

**Figure S1.** Covariate balance before and after matching among movers (A) who moved from high to low $PM_{2.5}$ areas and their eligible controls, (B) who moved from high to intermediate $PM_{2.5}$ areas and their eligible controls, (C) who moved from low to high $PM_{2.5}$ areas and their eligible controls, and (D) who moved from low to intermediate $PM_{2.5}$ areas and their eligible controls in Canada, 1997 to 2016

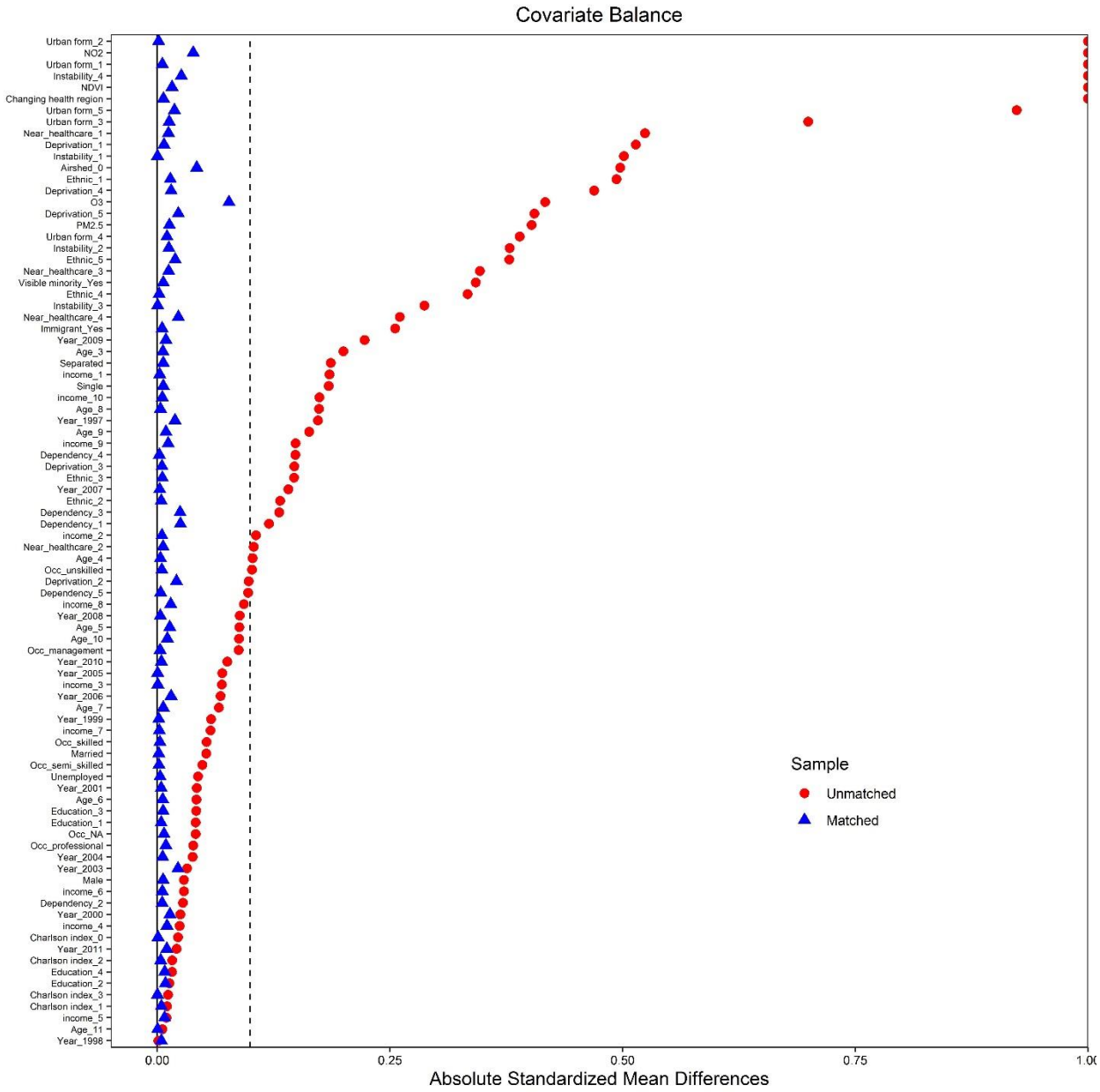
**Figure S2.** Covariate balance between movers from low to high $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

**Figure S3.** Covariate balance between movers from low to intermediate $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

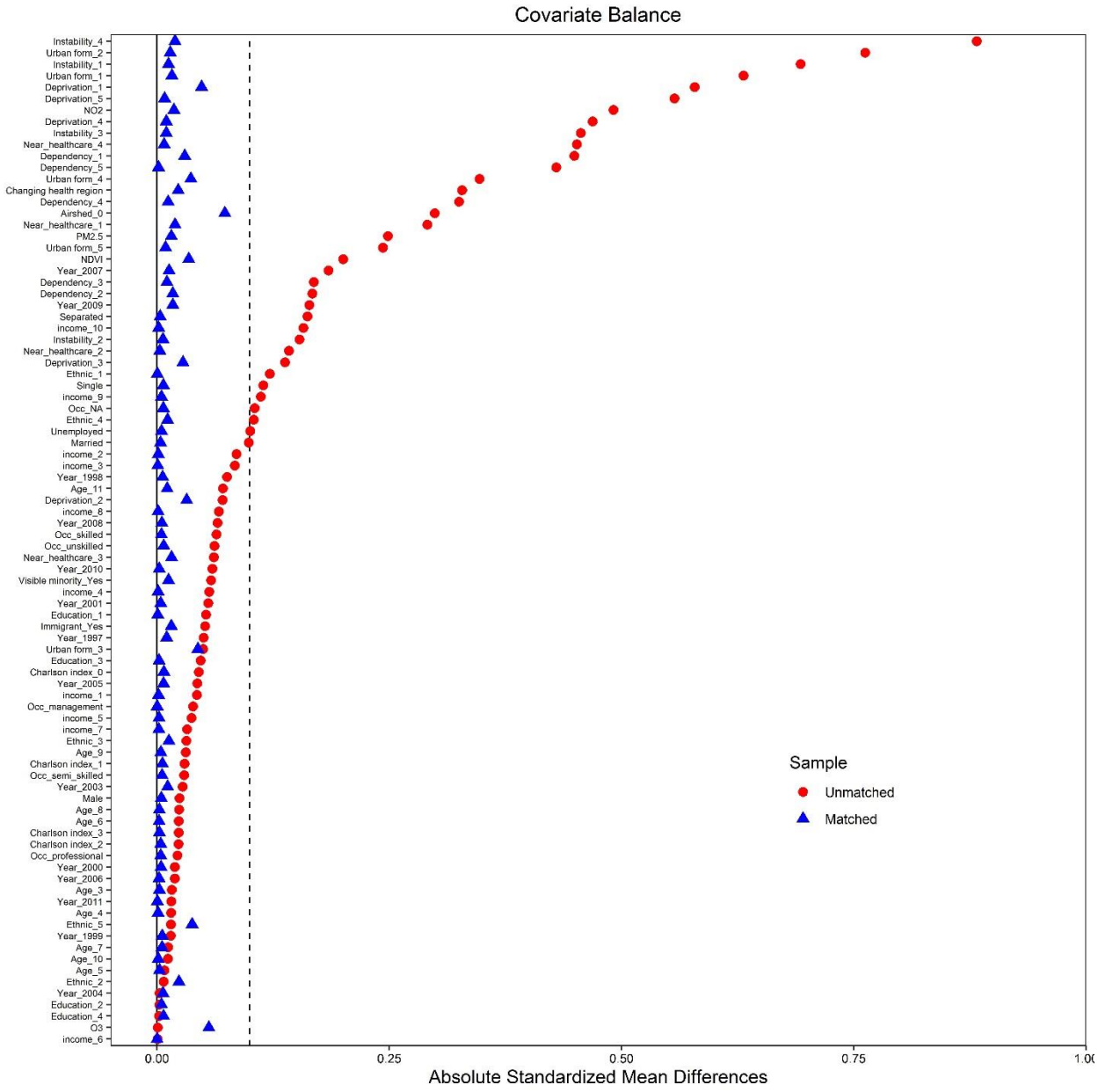
**Figure S4.** Covariate balance between movers from high to low $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

**Figure S5.** Covariate balance between movers from high to intermediate $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

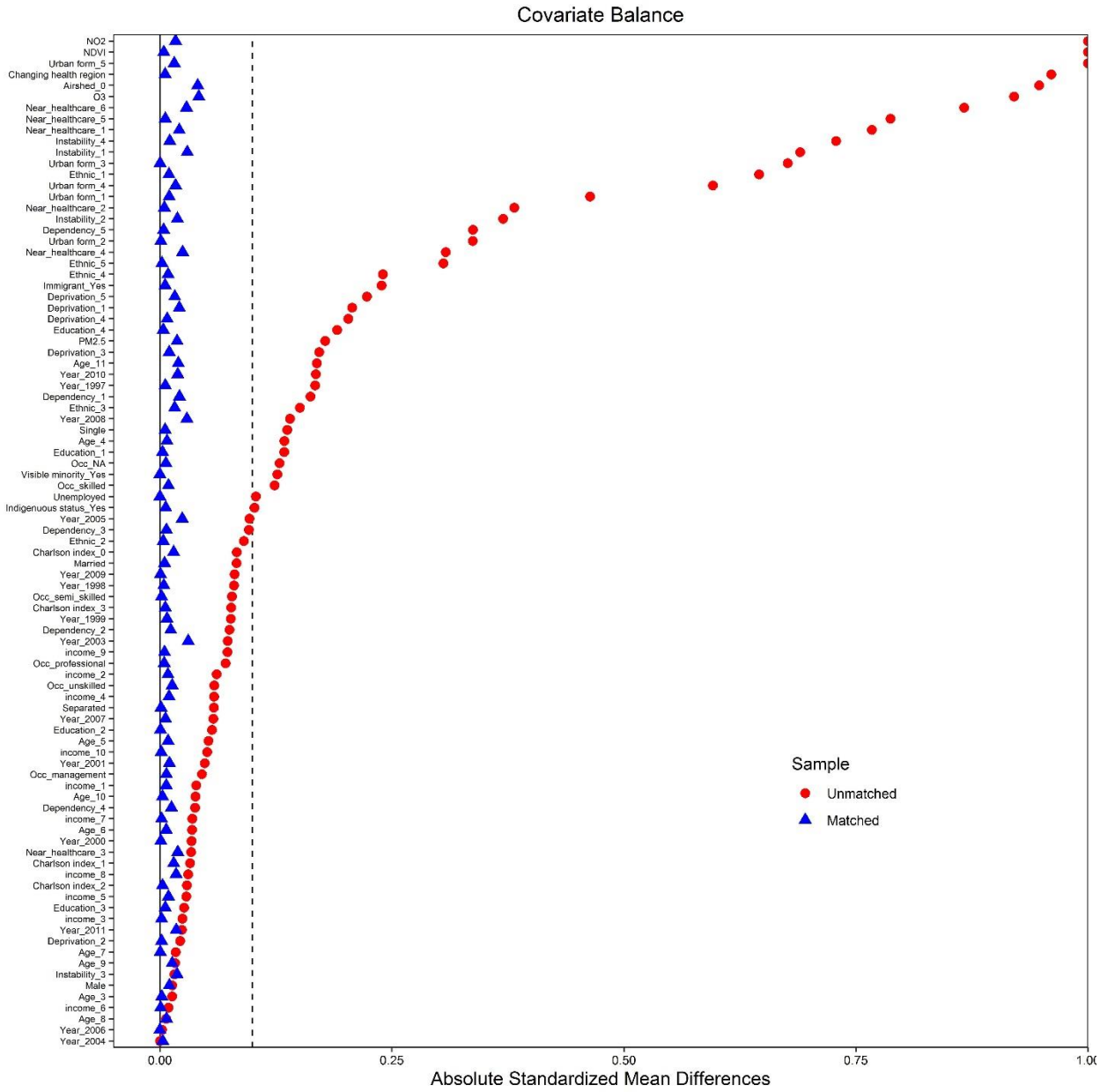
**Figure S6.** Distributions of residential exposure to $PM_{2.5}$ in movers who (A) moved from high to intermediate or low $PM_{2.5}$ areas (referred to as the treated) and their matched controls, as well as in movers who (B) moved from low to intermediate or high $PM_{2.5}$ areas and their matched controls, in Canada, 1997 to 2016. For each panel, the upper plot depicts $PM_{2.5}$ exposure before moving, the middle plot depicts $PM_{2.5}$ exposure after moving to intermediate $PM_{2.5}$ areas, and the lower plot depicts $PM_{2.5}$ exposure after moving from either high to low $PM_{2.5}$ or from low to high $PM_{2.5}$ areas


Covariate Balance
Urban form_2
NO2
Urban form_1
Instability_4
NDVI
Changing health region
Urban form_5
Urban form_3
Near_healthcare_1
Deprivation_1
Instability_1
Airshed_0
Ethnic_1
Deprivation_4
O3
Deprivation_5
PM2.5
Urban form_4
Instability_2
Ethnic_5
Near_healthcare_3
Visible minority_Yes
Ethnic_4
Instability_3
Near_healthcare_4
Immigrant_Yes
Year_2009
Age_3
Separated
income_1
Single
income_10
Age_8
Year_1997
Age_9
income_9
Dependency_4
Deprivation_3
Ethnic_3
Year_2007
Ethnic_2
Dependency_3
Dependency_1
income_2
Near_healthcare_2
Age_4
Occ_unskilled
Deprivation_2
Dependency_5
income_8
Year_2008
Age_5
Age_10
Occ_management
Year_2010
Year_2005
income_3
Year_2006
Age_7
Year_1999
income_7
Occ_skilled
Married
Occ_semi_skilled
Unemployed
Year_2001
Age_6
Education_3
Education_1
Occ_NA
Occ_professional
Year_2004
Year_2003
Male
income_6
Dependency_2
Year_2000
income_4
Charlson index_0
Year_2011
Charlson index_2
Education_4
Education_2
Charlson index_3
Charlson index_1
income_5
Age_11
Year_1998
Sample
Unmatched
Matched
0.00
0.25
0.50
0.75
1.00
Absolute Standardized Mean Differences

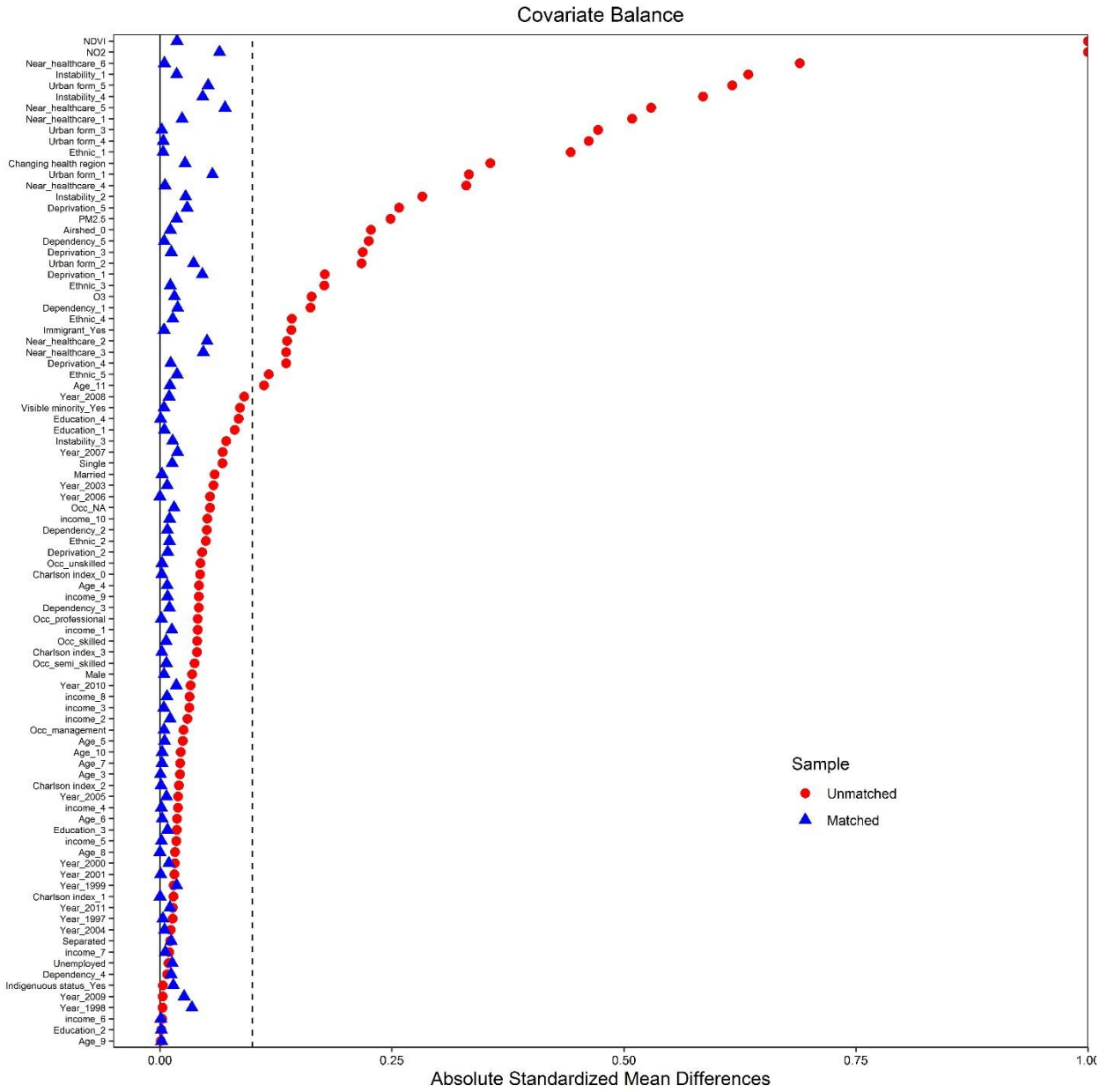
**(A)**


Covariate Balance
Instability_4
Urban form_2
Instability_1
Urban form_1
Deprivation_1
Deprivation_5
NO2
Deprivation_4
Instability_3
Near_healthcare_4
Dependency_1
Dependency_5
Urban form_4
Changing health region
Dependency_4
Airshed_0
Near_healthcare_1
PM2.5
Urban form_5
NDVI
Year_2007
Dependency_3
Dependency_2
Year_2009
Separated
income_10
Instability_2
Near_healthcare_2
Deprivation_3
Ethnic_1
Single
income_9
Occ_NA
Ethnic_4
Unemployed
Married
income_2
income_3
Year_1998
Age_11
Deprivation_2
income_8
Year_2008
Occ_skilled
Occ_unskilled
Near_healthcare_3
Year_2010
Visible minority_Yes
income_4
Year_2001
Education_1
Immigrant_Yes
Year_1997
Urban form_3
Education_3
Charlson index_0
Year_2005
income_1
Occ_management
income_5
income_7
Ethnic_3
Age_9
Charlson index_1
Occ_semi_skilled
Year_2003
Male
Age_8
Age_6
Charlson index_3
Charlson index_2
Occ_professional
Year_2000
Year_2006
Age_3
Year_2011
Age_4
Ethnic_5
Year_1999
Age_7
Age_10
Age_5
Ethnic_2
Year_2004
Education_2
Education_4
O3
income_6
Sample
Unmatched
Matched
0.00
0.25
0.50
0.75
1.00
Absolute Standardized Mean Differences

**(B)**

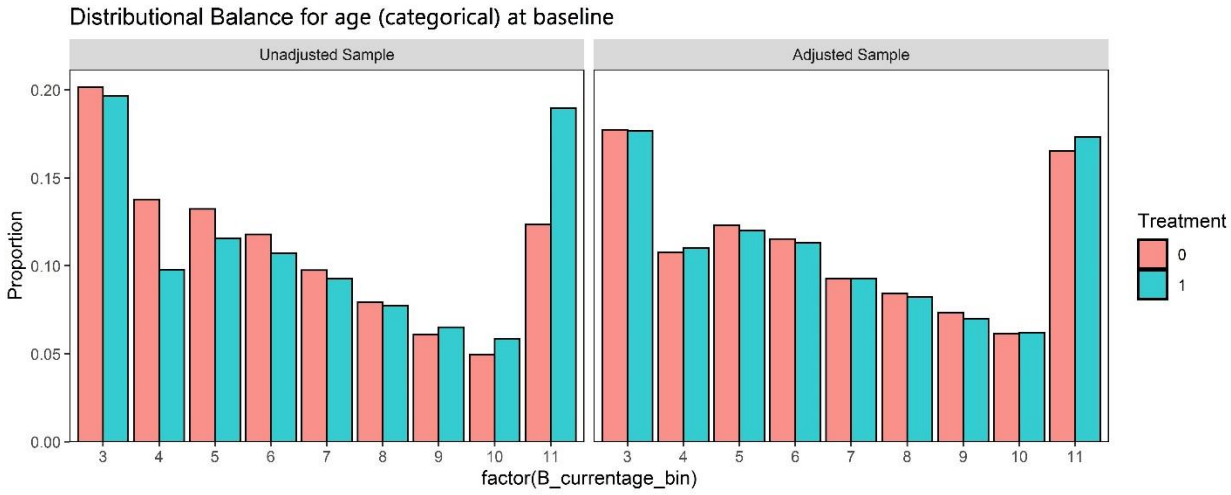

Covariate Balance
NO2
NDVI
Urban form_5
Changing health region
Airshed_0
O3
Near_healthcare_6
Near_healthcare_5
Near_healthcare_1
Instability_4
Instability_1
Urban form_3
Ethnic_1
Urban form_4
Urban form_1
Near_healthcare_2
Instability_2
Dependency_5
Urban form_2
Near_healthcare_4
Ethnic_5
Ethnic_4
Immigrant_Yes
Deprivation_5
Deprivation_1
Deprivation_4
Education_4
PM2.5
Deprivation_3
Age_11
Year_2010
Year_1997
Dependency_1
Ethnic_3
Year_2008
Single
Age_4
Education_1
Occ_NA
Visible minority_Yes
Occ_skilled
Unemployed
Indigenuous status_Yes
Year_2005
Dependency_3
Ethnic_2
Charlson index_0
Married
Year_2009
Year_1998
Occ_semi_skilled
Charlson index_3
Year_1999
Dependency_2
Year_2003
income_9
Occ_professional
income_2
Occ_unskilled
income_4
Separated
Year_2007
Education_2
Age_5
income_10
Year_2001
Occ_management
income_1
Age_10
Dependency_4
income_7
Age_6
Year_2000
Near_healthcare_3
Charlson index_1
income_8
Charlson index_2
income_5
Education_3
income_3
Year_2011
Deprivation_2
Age_7
Age_9
Instability_3
Male
Age_3
income_6
Age_8
Year_2006
Year_2004
Sample
Unmatched
Matched
0.00
0.25
0.50
0.75
1.00
Absolute Standardized Mean Differences

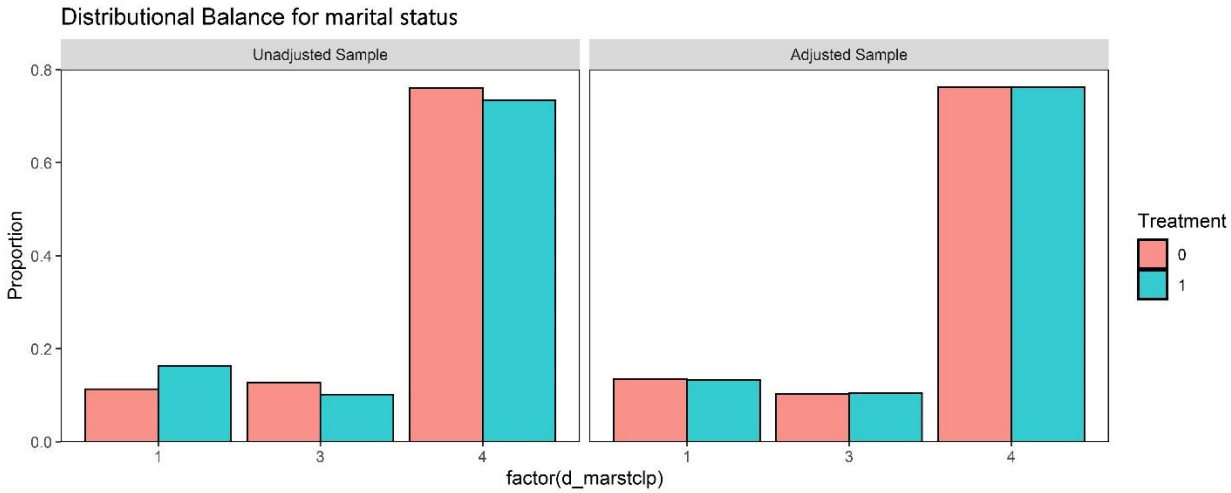
**(C)**

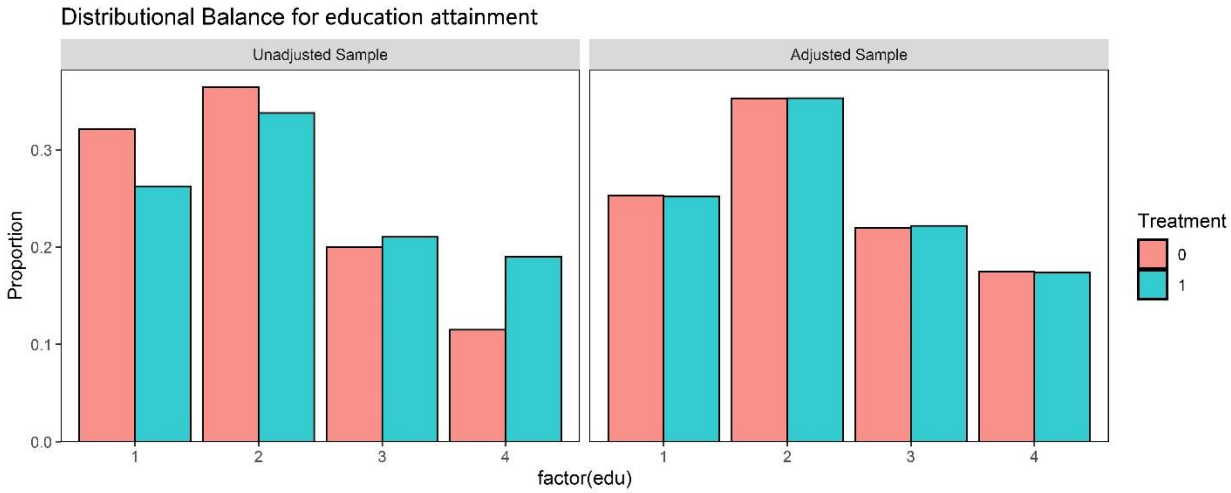


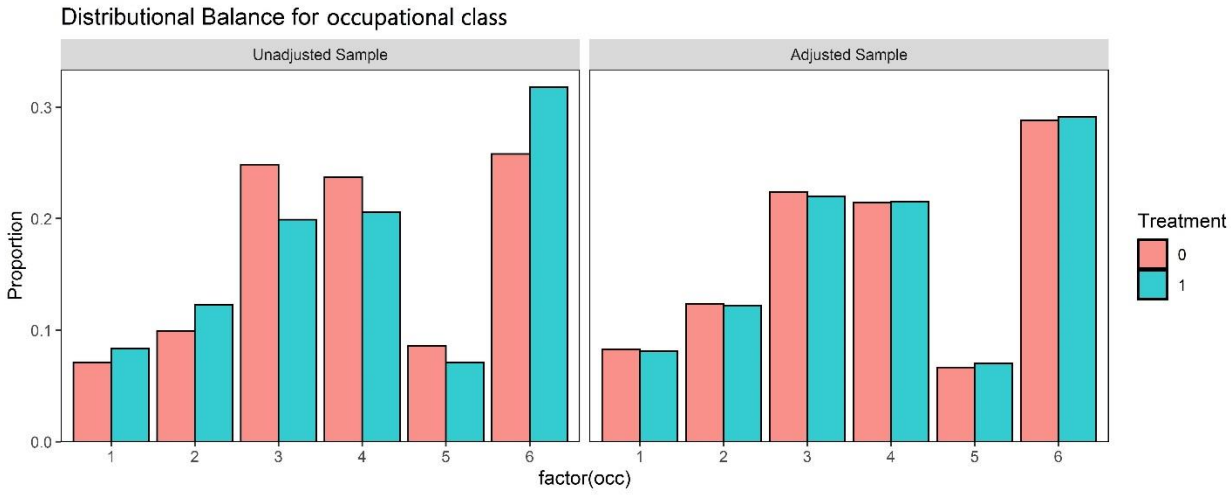

**(D)**

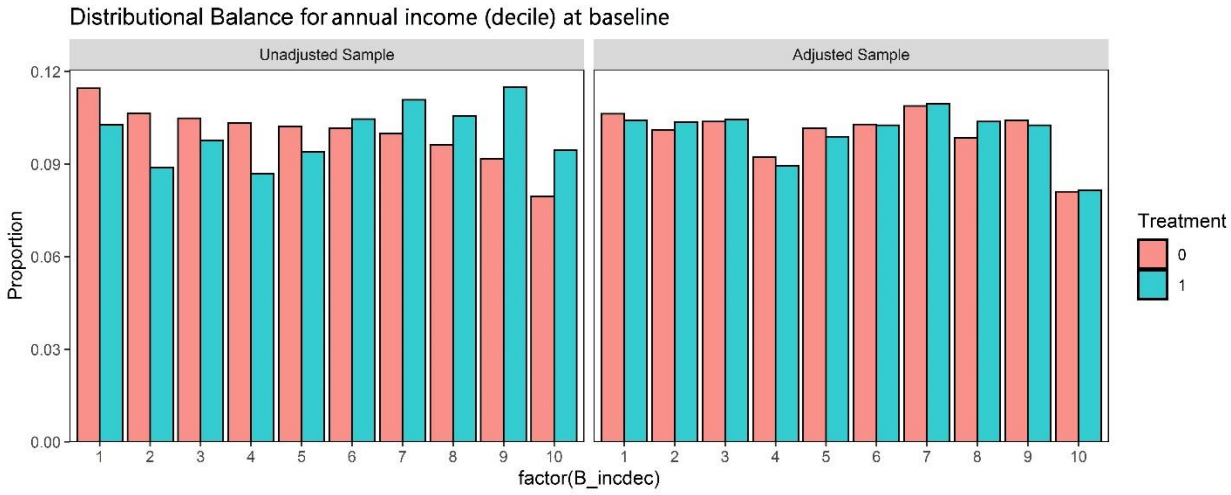
**Figure S1.** Covariate balance before and after matching among movers (A) who moved from high to low $PM_{2.5}$ areas and their eligible controls, (B) who moved from high to intermediate $PM_{2.5}$ areas and their eligible controls, (C) who moved from low to high $PM_{2.5}$ areas and their eligible controls, and (D) who moved from low to intermediate $PM_{2.5}$ areas and their eligible controls in Canada, 1997 to 2016

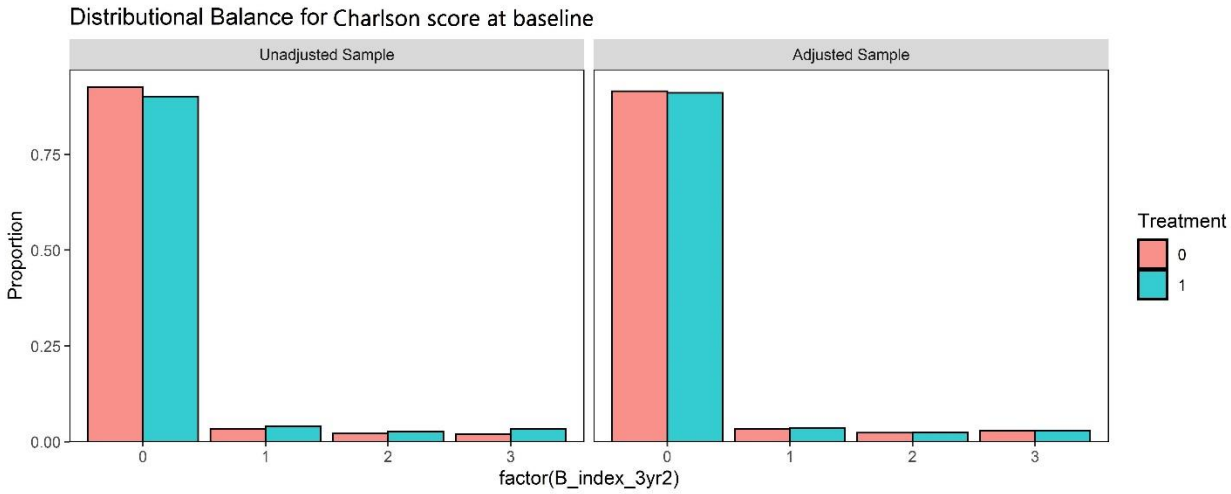

Distributional Balance for age (categorical) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.05
0.10
0.15
0.20
3
4
5
6
7
8
9
10
11
factor(B_currentage_bin)
Treatment
0
1

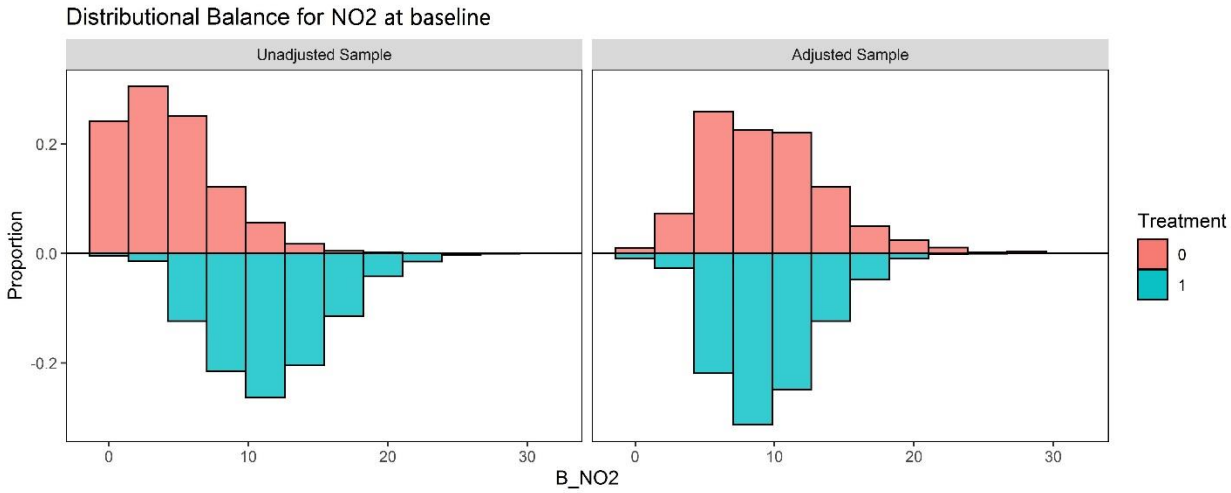

Distributional Balance for marital status
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0.8
1
3
4
factor(d_marstclp)
Treatment
0
1

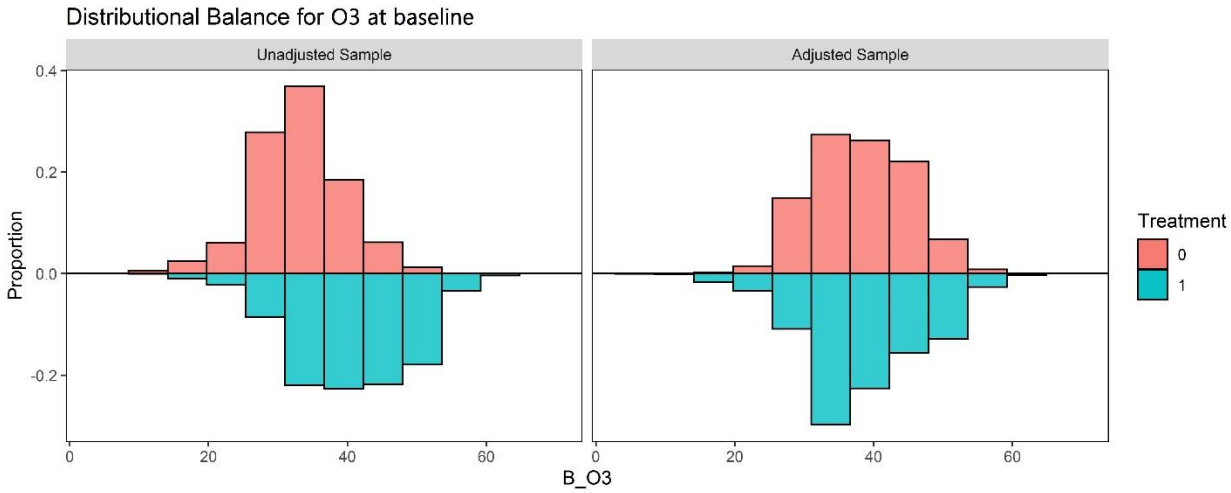

Distributional Balance for education attainment
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
1
2
3
4
factor(edu)
Treatment
0
1

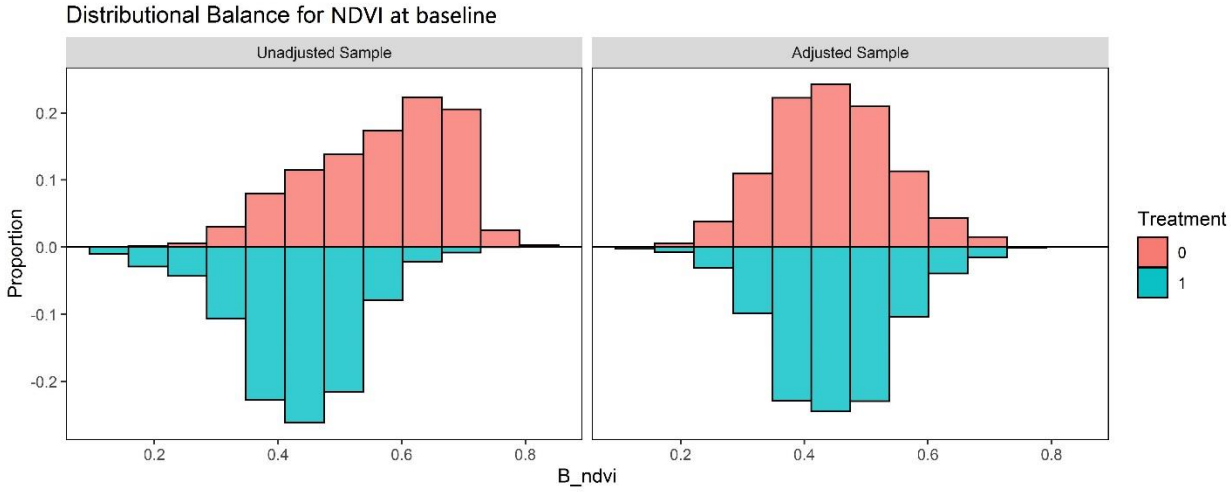

Distributional Balance for occupational class
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
1
2
3
4
5
6
factor(occ)
Treatment
0
1

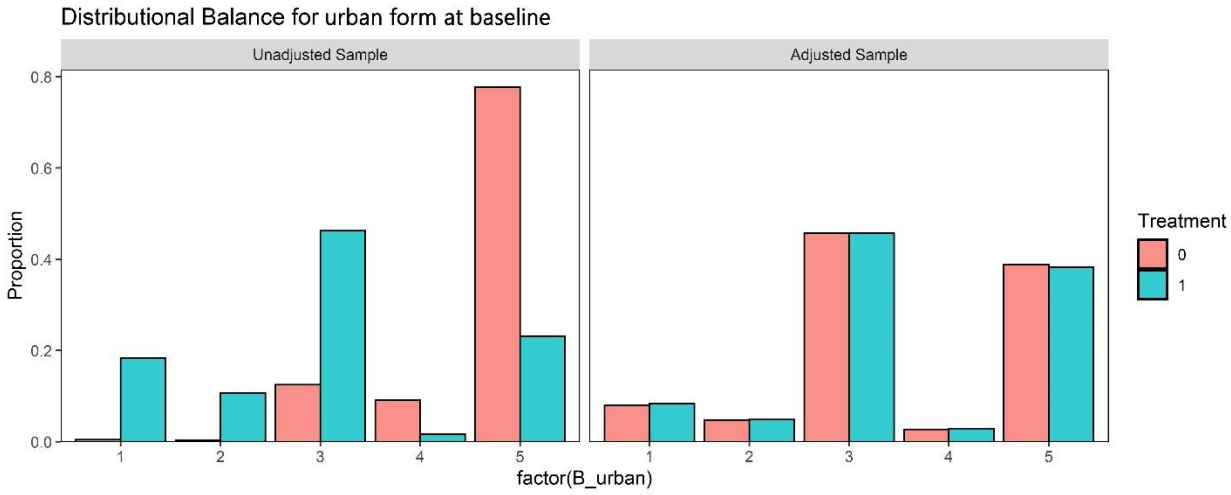

Distributional Balance for annual income (decile) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.03
0.06
0.09
0.12
1
2
3
4
5
6
7
8
9
10
factor(B_incdec)
Treatment
0
1

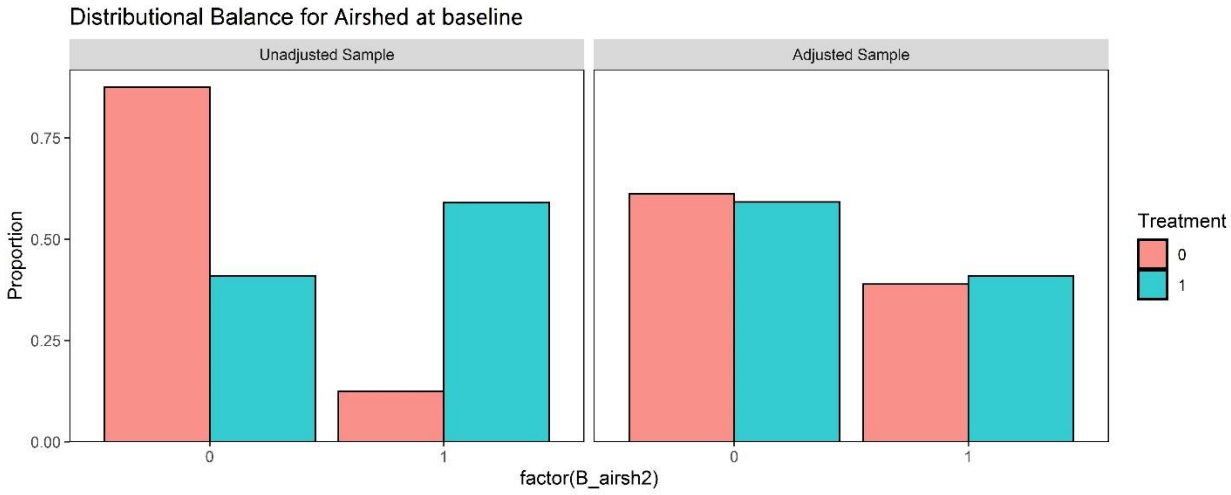

Distributional Balance for Charlson score at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
2
3
factor(B_index_3yr2)
Treatment
0
1

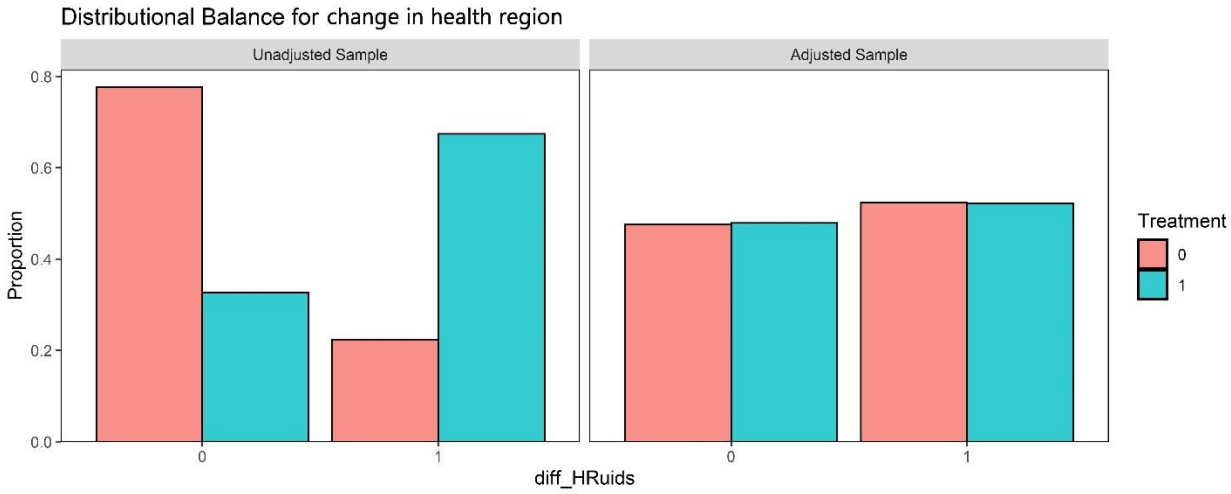

Distributional Balance for NO2 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
0
10
20
30
B_NO2
Treatment
0
1

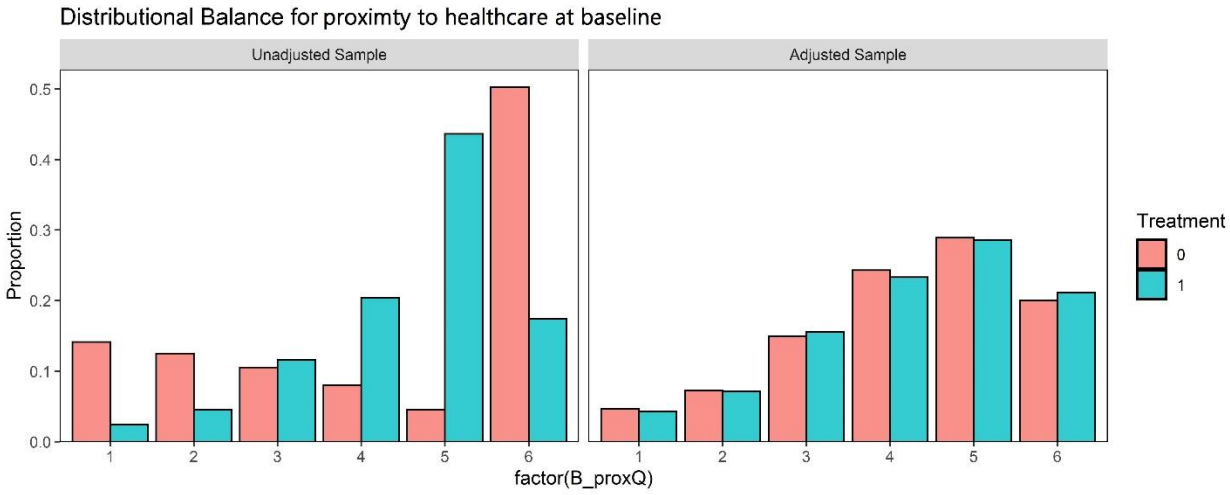

Distributional Balance for O3 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.4
0.2
0.0
-0.2
0
20
40
60
B_O3
Treatment
0
1

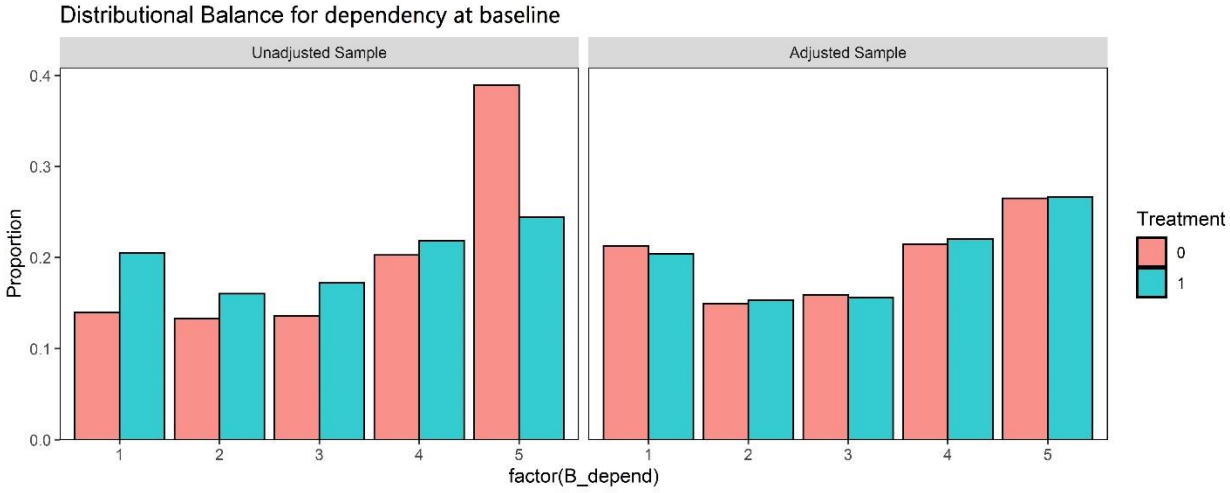

Distributional Balance for NDVI at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.1
0.0
-0.1
-0.2
0.2
0.4
0.6
0.8
B_ndvi
Treatment
0
1

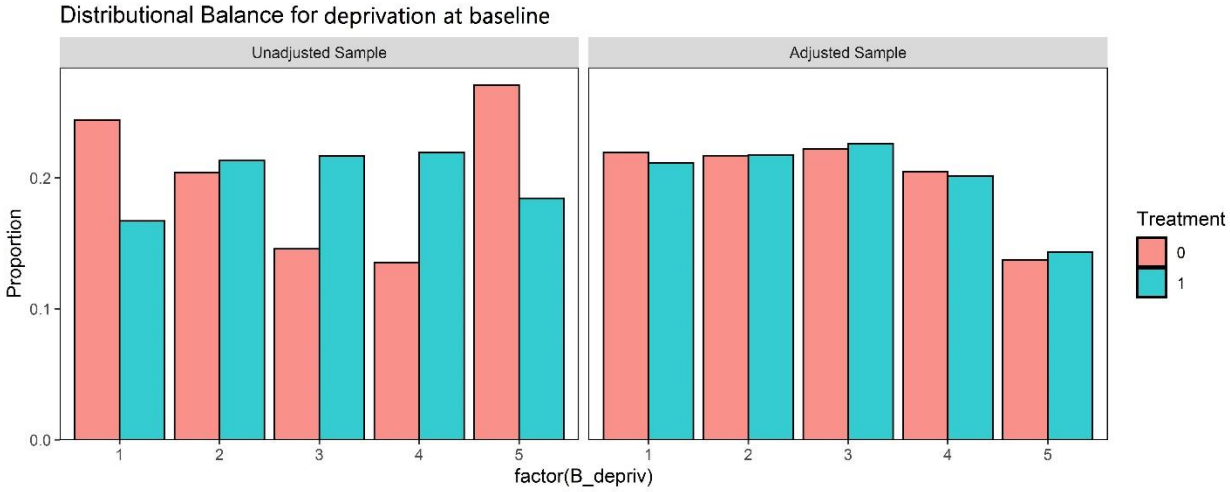

Distributional Balance for urban form at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0.8
1
2
3
4
5
factor(B_urban)
Treatment
0
1

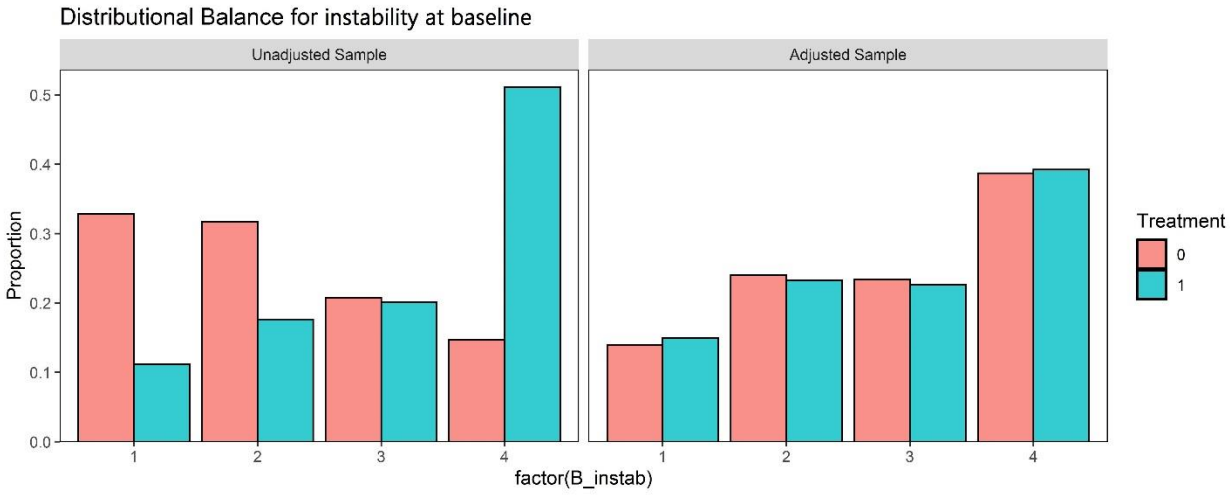

Distributional Balance for Airshed at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
factor(B_airsh2)
Treatment
0
1

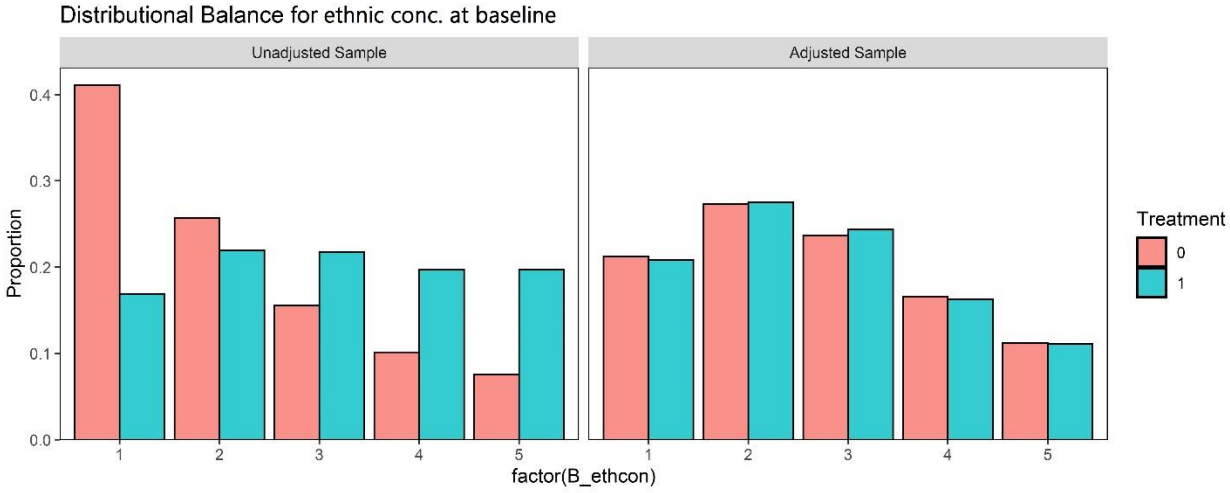

Distributional Balance for change in health region
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0.8
0
1
diff_HRuids
Treatment
0
1

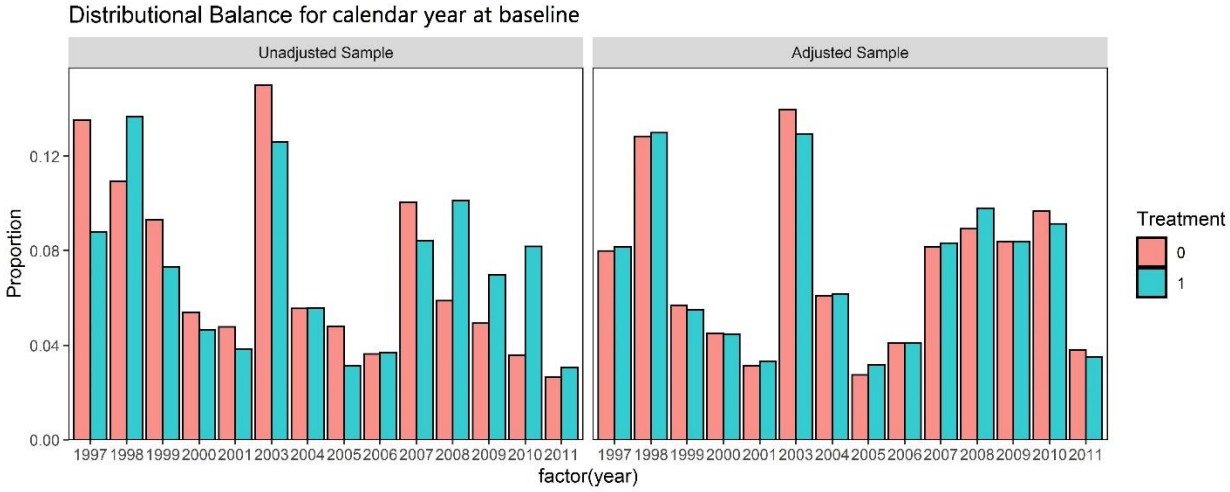

Distributional Balance for proximty to healthcare at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
factor(B_proxQ)
Treatment
0
1


Distributional Balance for dependency at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_depend)
Treatment
0
1


Distributional Balance for deprivation at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
factor(B_depriv)
Treatment
0
1


Distributional Balance for instability at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
factor(B_instab)
Treatment
0
1


Distributional Balance for ethnic conc. at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_ethcon)
Treatment
0
1


Distributional Balance for calendar year at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.04
0.08
0.12
1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
factor(year)
Treatment
0
1

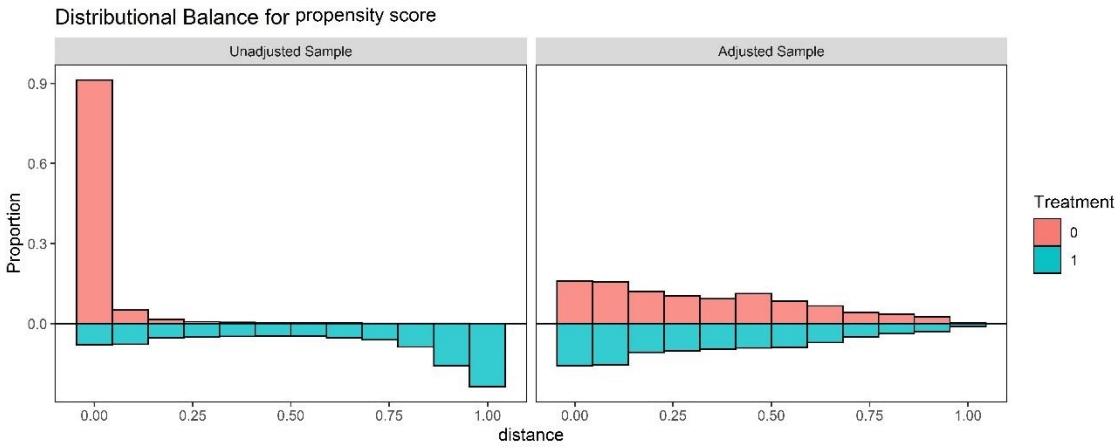
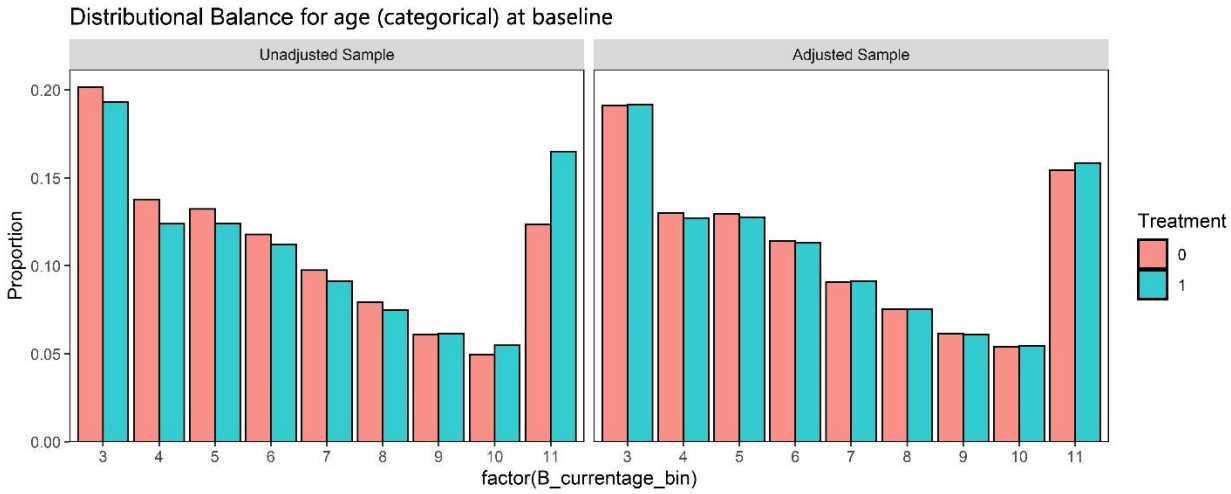
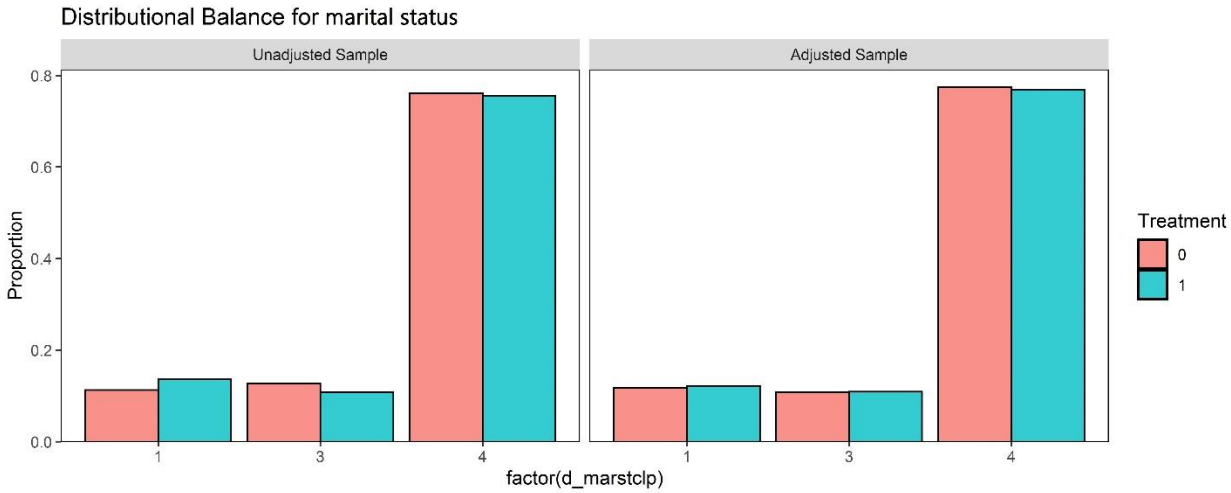
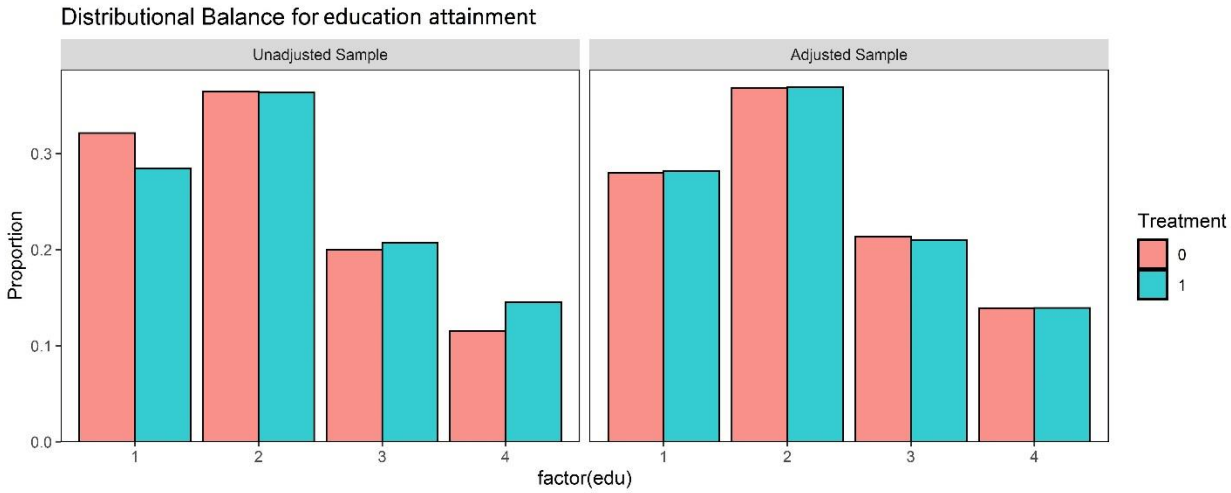
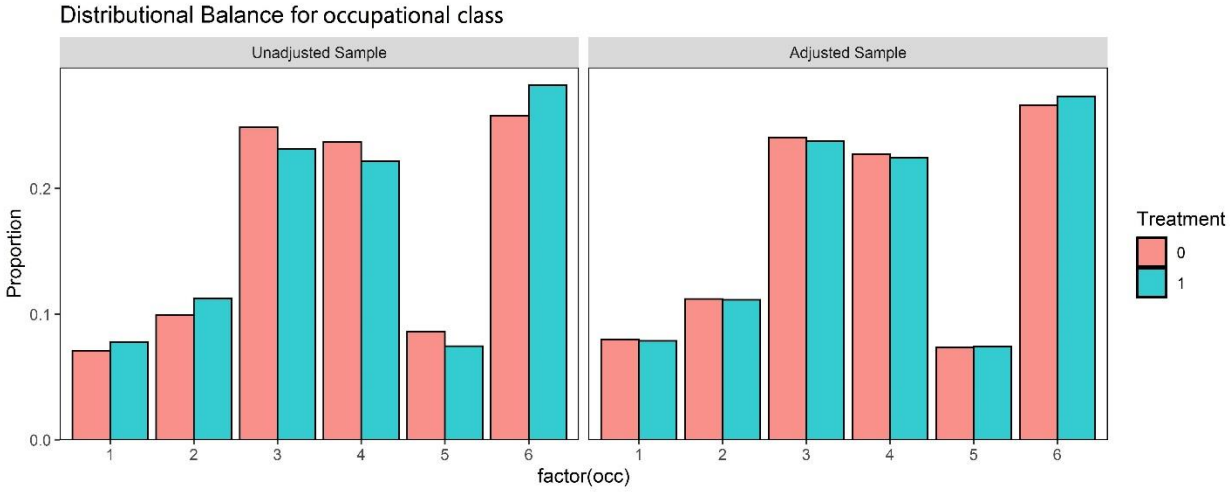


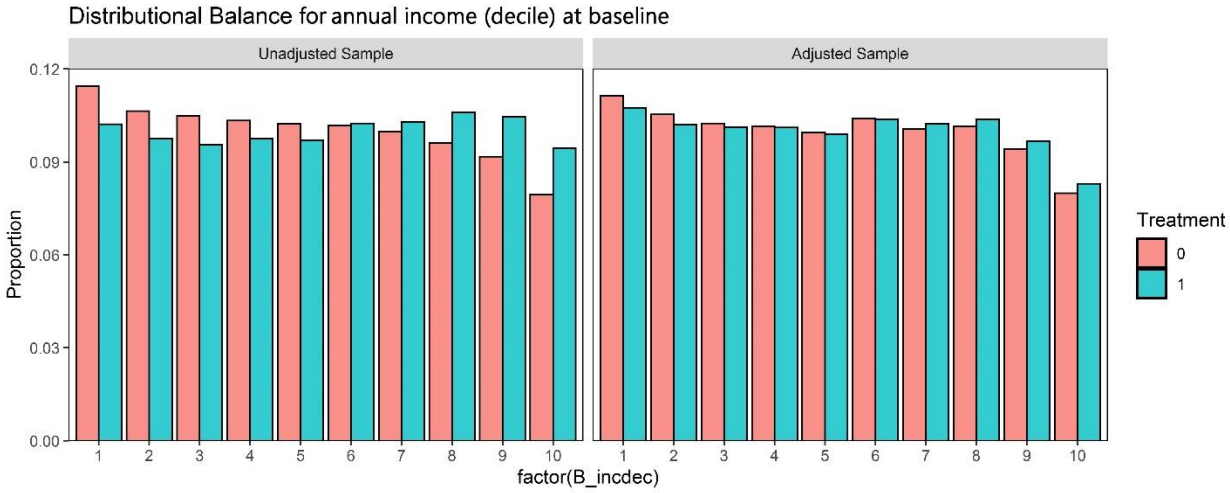

**Figure S2.** Covariate balance between movers from low to high $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

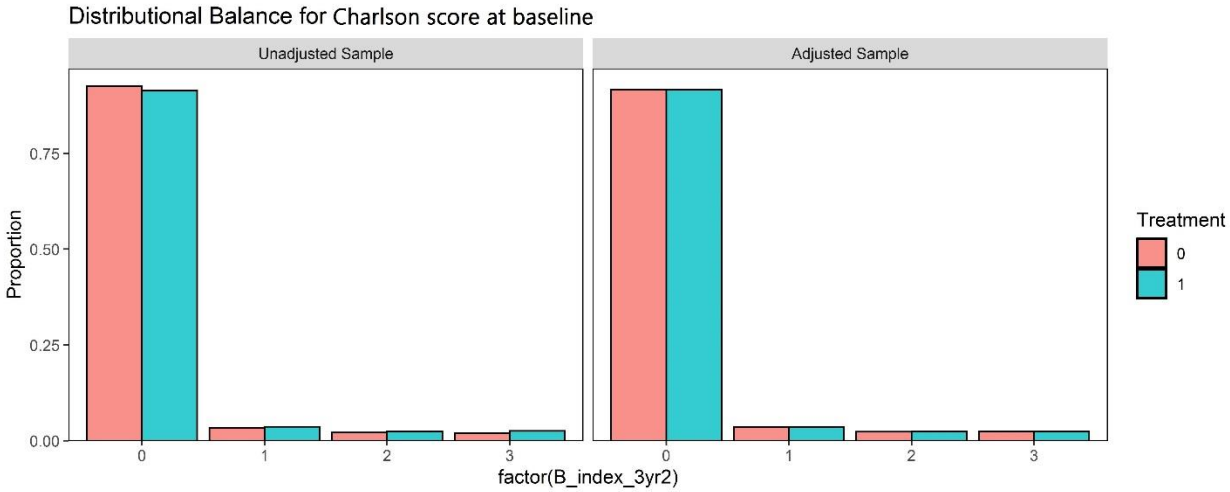

Distributional Balance for age (categorical) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.20
0.15
0.10
0.05
0.00
3 4 5 6 7 8 9 10 11
factor(B_currentage_bin)
Treatment
0
1

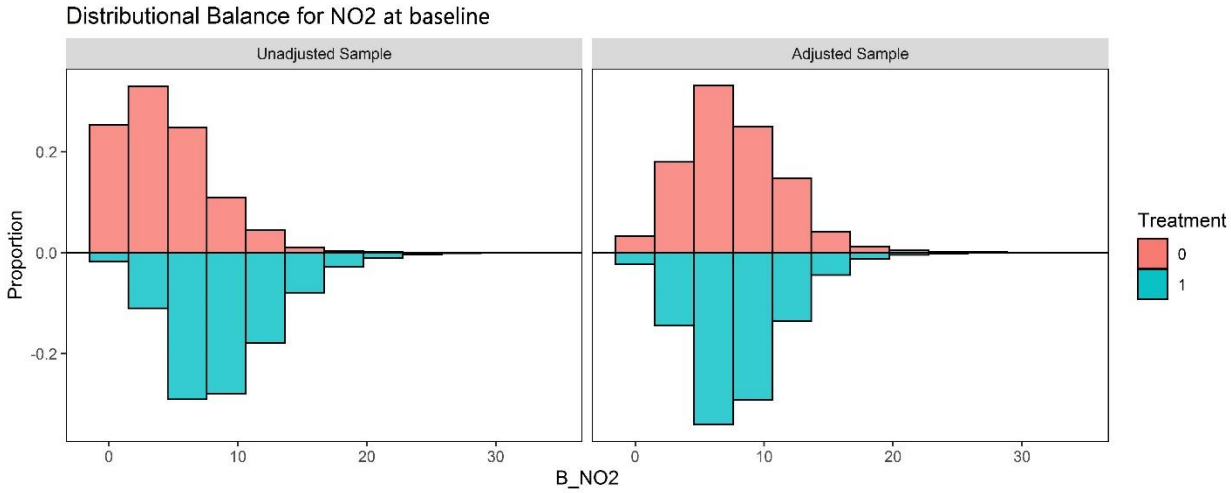

Distributional Balance for marital status
Unadjusted Sample
Adjusted Sample
Proportion
0.8
0.6
0.4
0.2
0.0
1 3 4
factor(d_marstclp)
Treatment
0
1

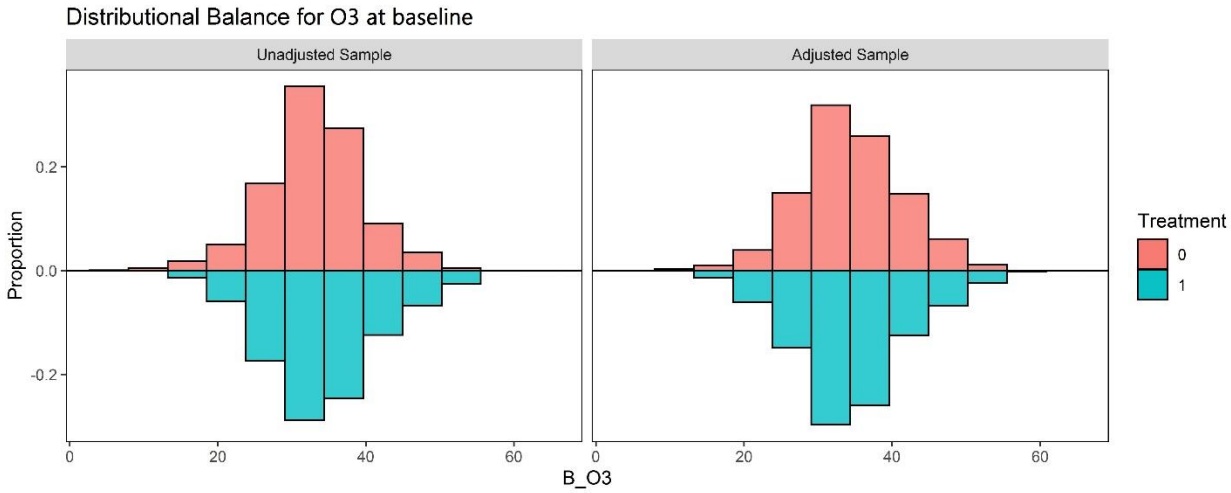

Distributional Balance for education attainment
Unadjusted Sample
Adjusted Sample
Proportion
0.3
0.2
0.1
0.0
1 2 3 4
factor(edu)
Treatment
0
1

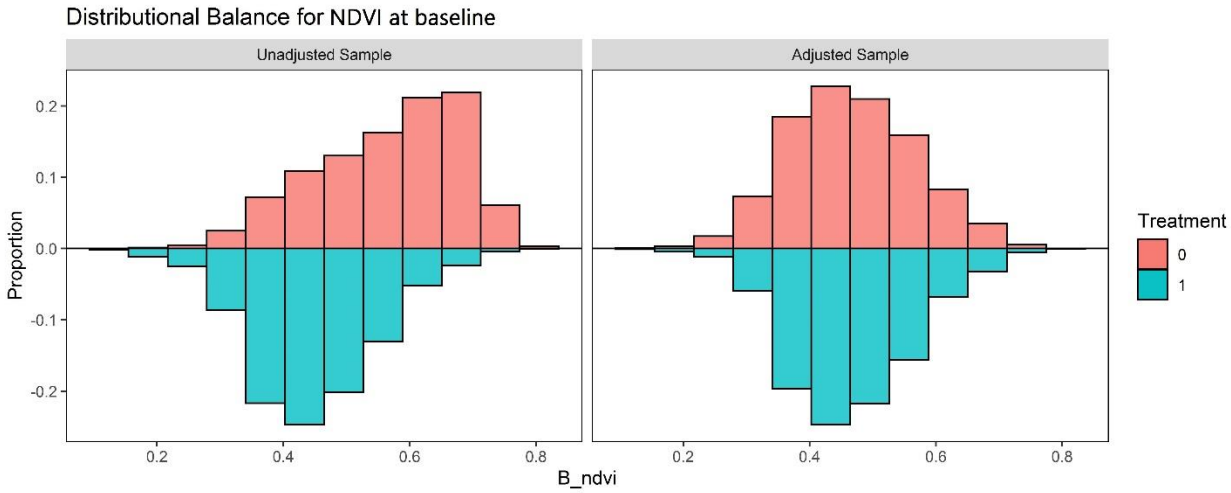

Distributional Balance for occupational class
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
6
factor(occ)
Treatment
0
1

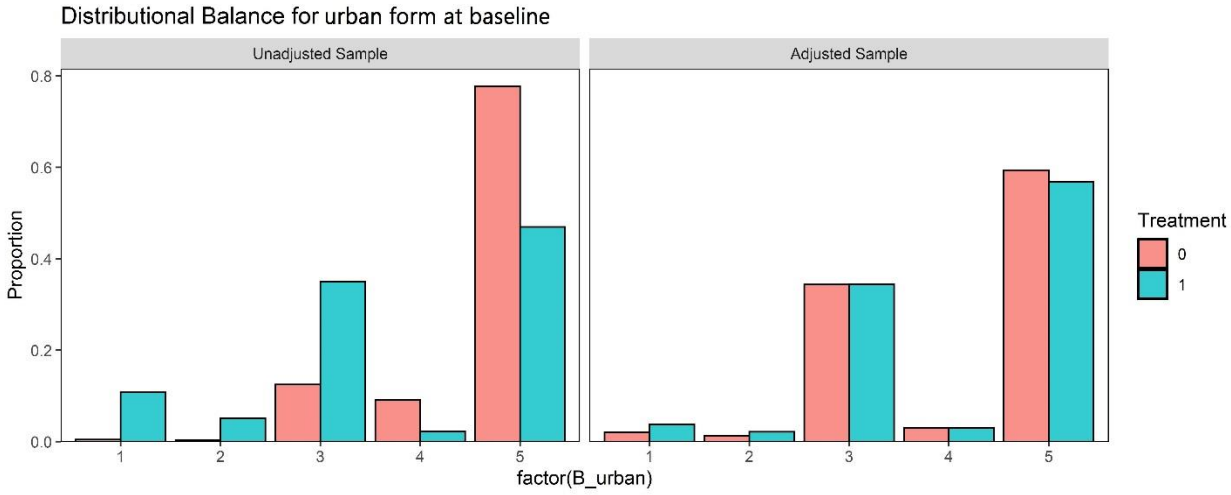

Distributional Balance for annual income (decile) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.03
0.06
0.09
0.12
1
2
3
4
5
6
7
8
9
10
factor(B_incdec)
Treatment
0
1

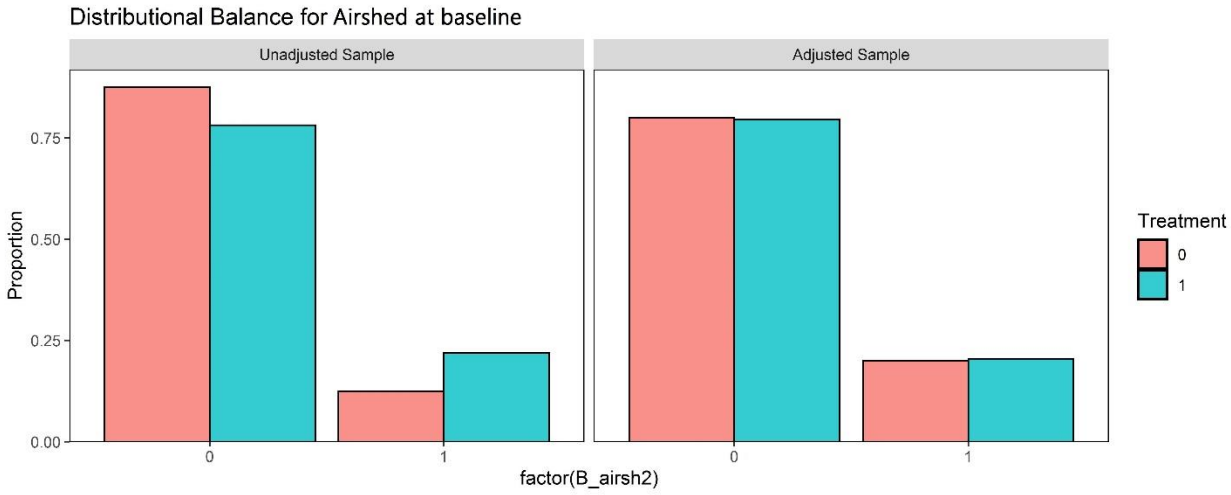

Distributional Balance for Charlson score at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
2
3
factor(B_index_3yr2)
Treatment
0
1

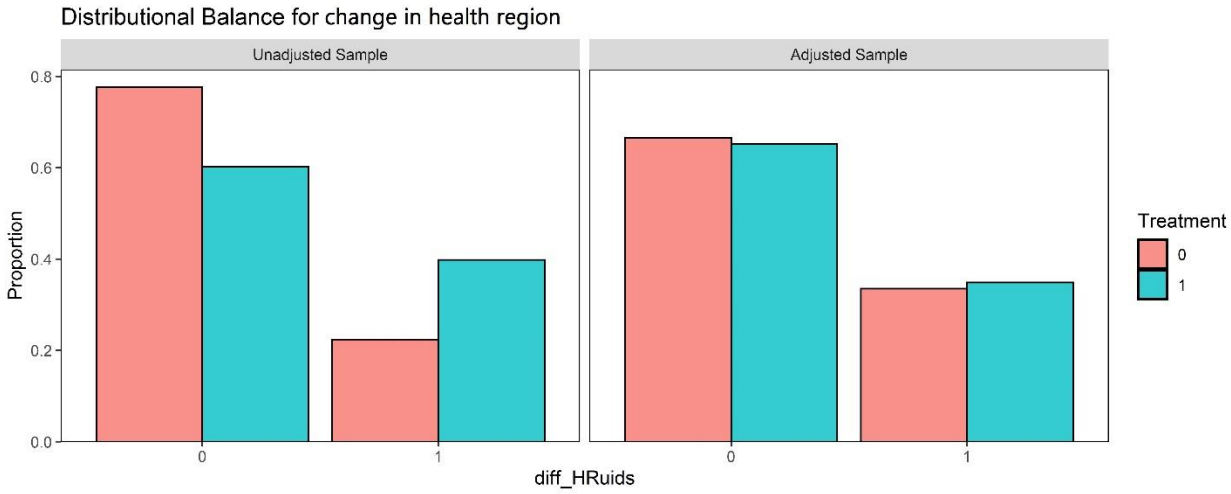

Distributional Balance for NO2 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
0
10
20
30
B_NO2
Treatment
0
1

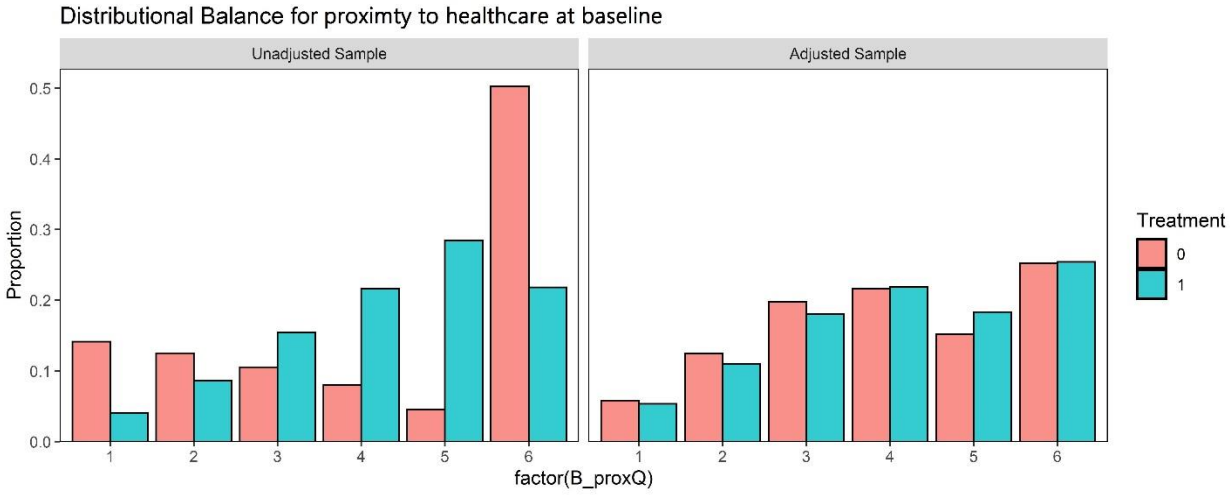

Distributional Balance for O3 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
0
20
40
60
B_O3
Treatment
0
1

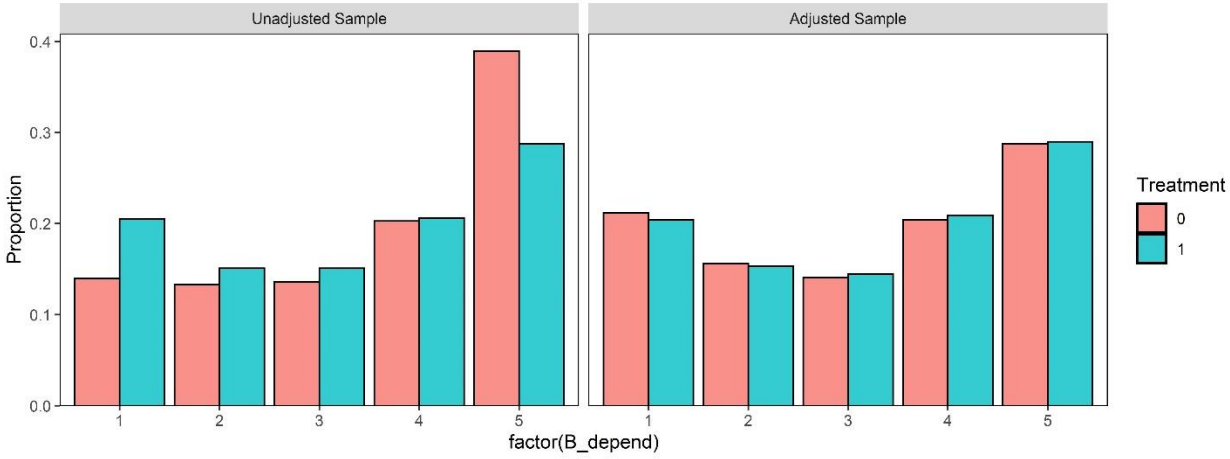

Distributional Balance for NDVI at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.1
0.0
-0.1
-0.2
0.2
0.4
0.6
0.8
B_ndvi
Treatment
0
1

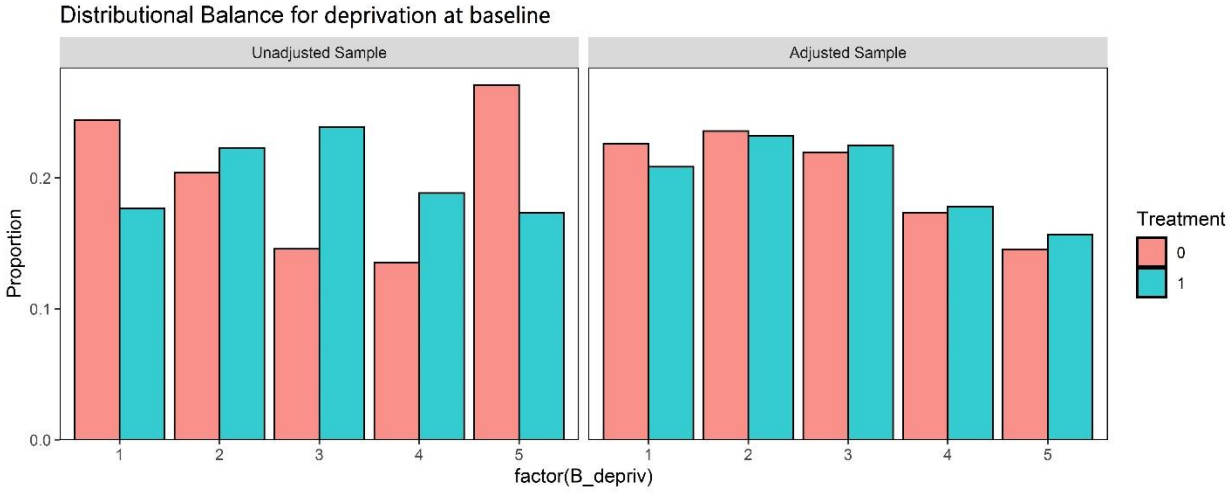

Distributional Balance for urban form at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0.8
1
2
3
4
5
factor(B_urban)
Treatment
0
1

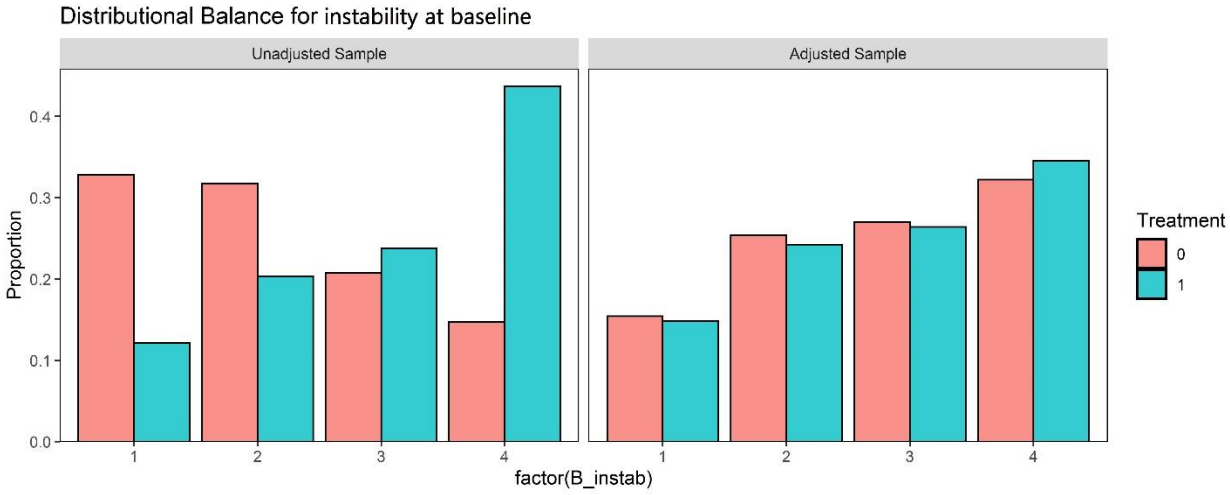

Distributional Balance for Airshed at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
factor(B_airsh2)
Treatment
0
1

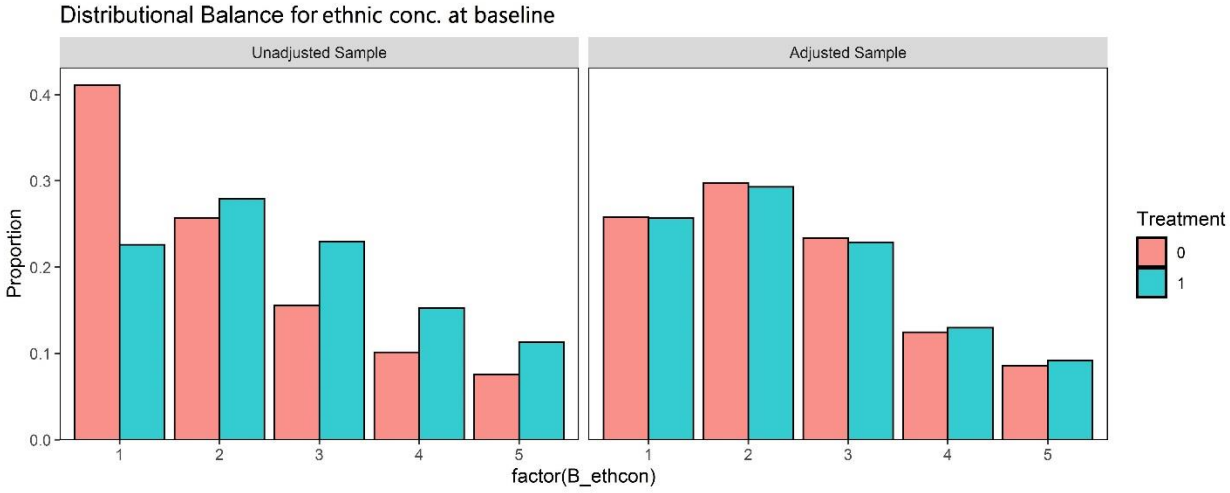

Distributional Balance for change in health region
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0.8
0
1
diff_HRuids
Treatment
0
1

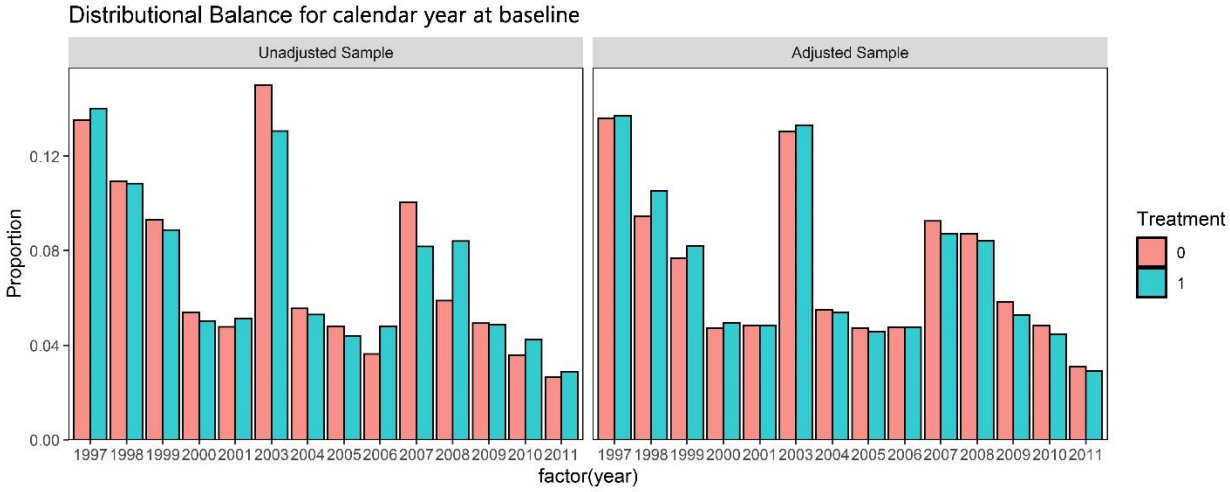

Distributional Balance for proximty to healthcare at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
factor(B_proxQ)
Treatment
0
1


Distributional Balance for dependency at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_depend)
Treatment
0
1


Distributional Balance for deprivation at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
factor(B_depriv)
Treatment
0
1


Distributional Balance for instability at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
factor(B_instab)
Treatment
0
1


Distributional Balance for ethnic conc. at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_ethcon)
Treatment
0
1


Distributional Balance for calendar year at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.04
0.08
0.12
1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
factor(year)
Treatment
0
1

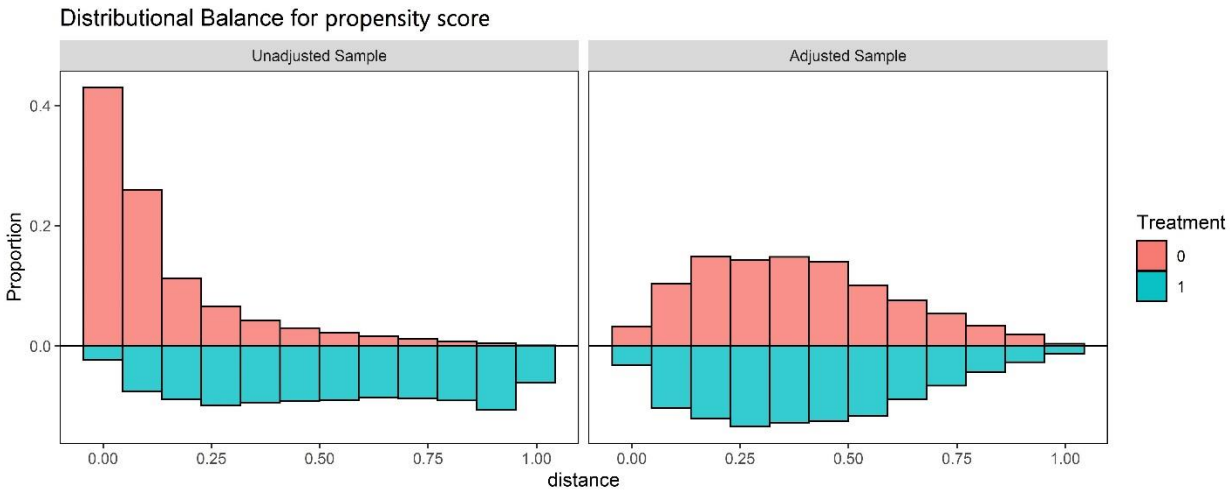
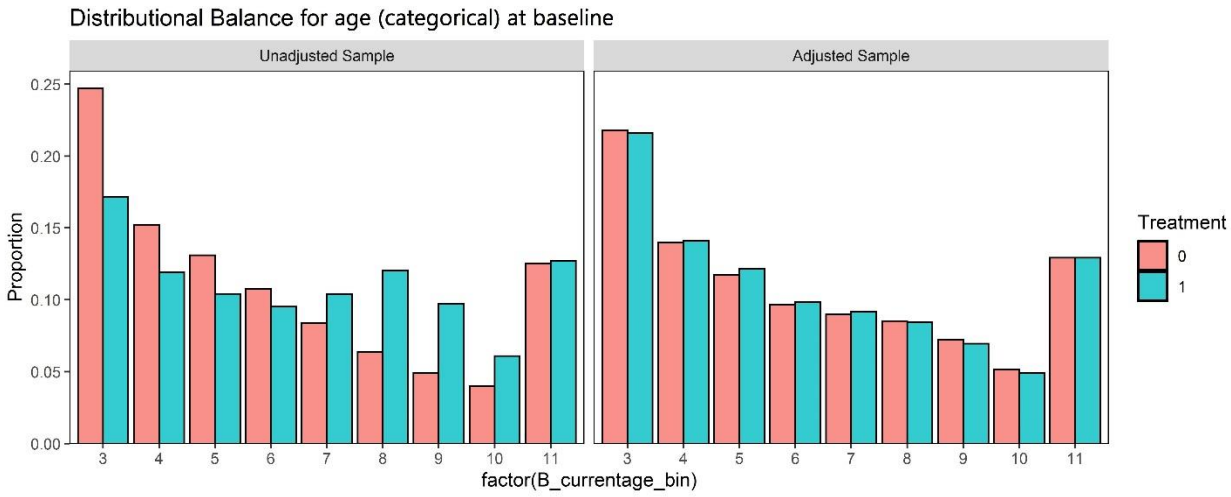
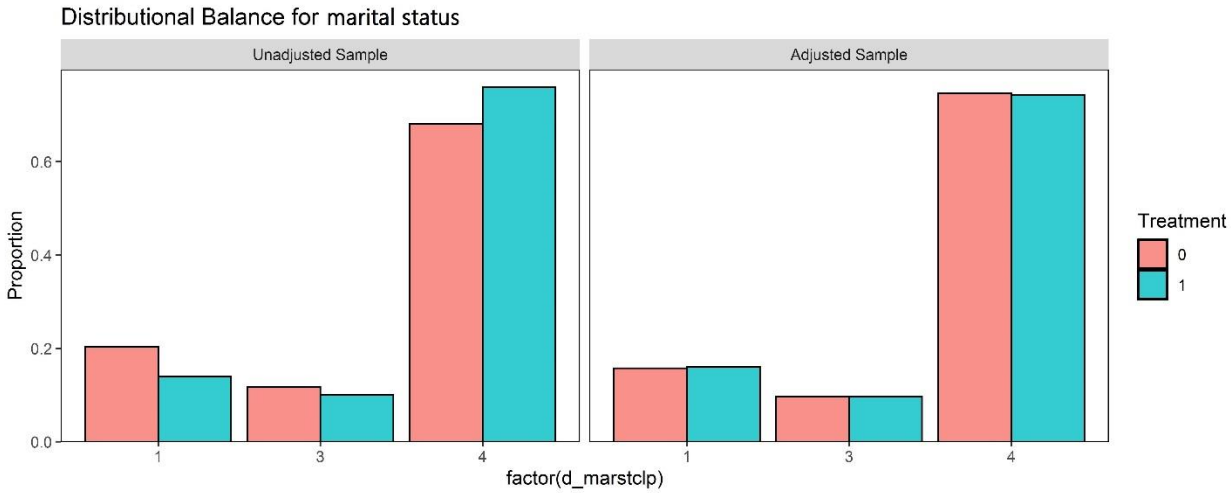
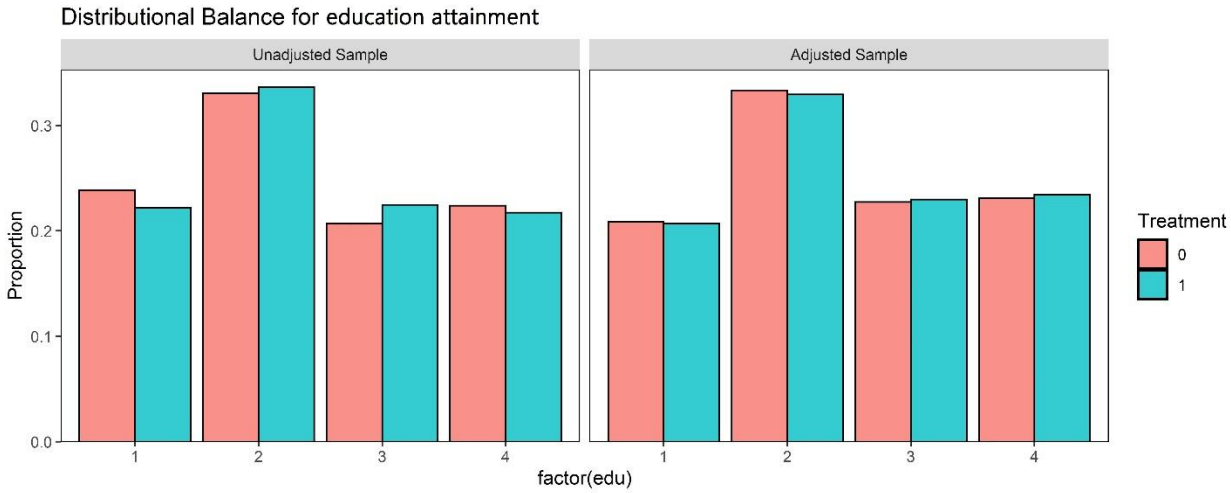
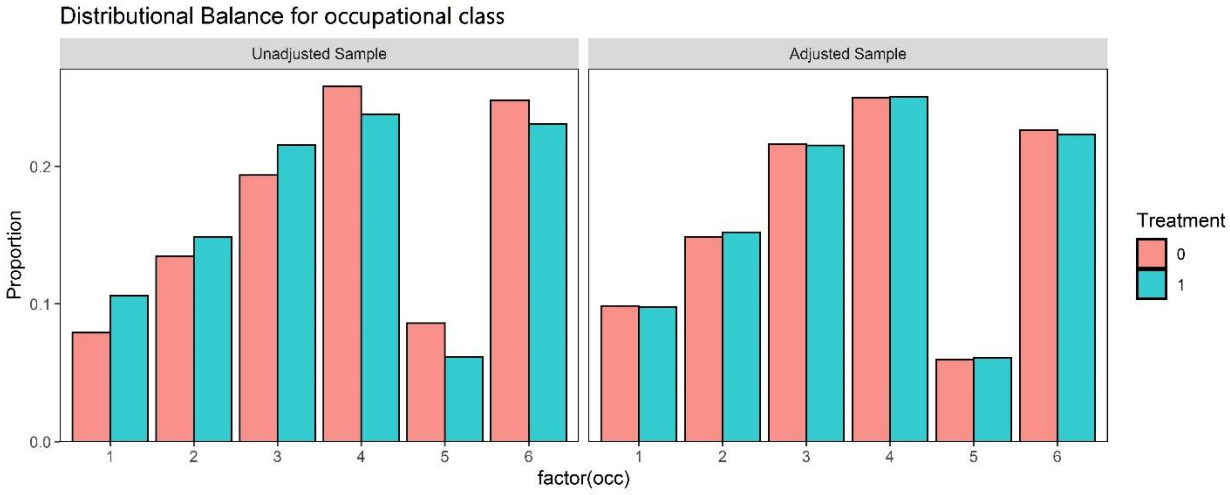


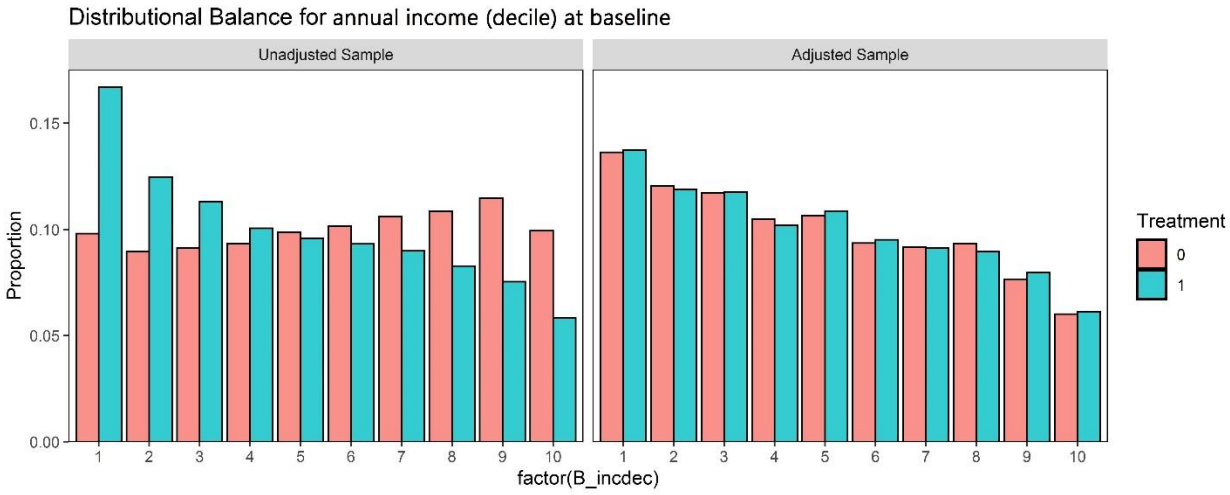

**Figure S3.** Covariate balance between movers from low to intermediate $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

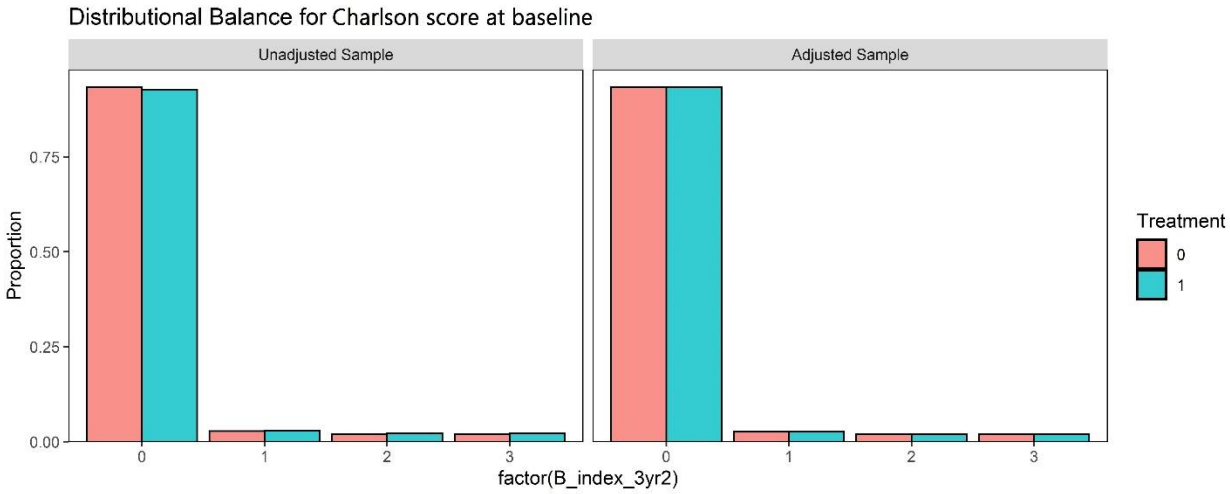

Distributional Balance for age (categorical) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.25
0.20
0.15
0.10
0.05
0.00
3 4 5 6 7 8 9 10 11
factor(B_currentage_bin)
Treatment
0
1

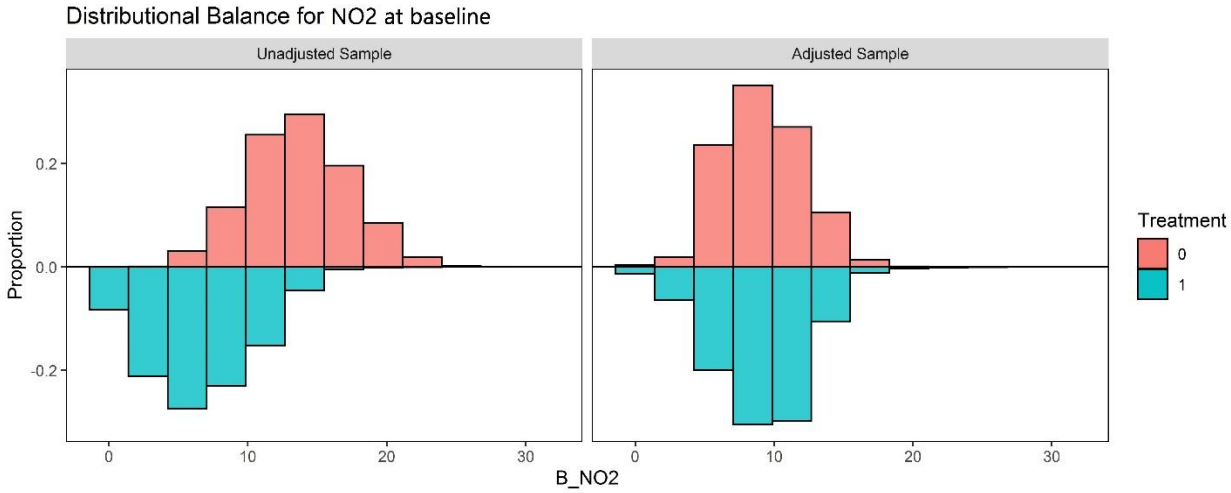

Distributional Balance for marital status
Unadjusted Sample
Adjusted Sample
Proportion
0.6
0.4
0.2
0.0
1 3 4
factor(d_marstclp)
Treatment
0
1

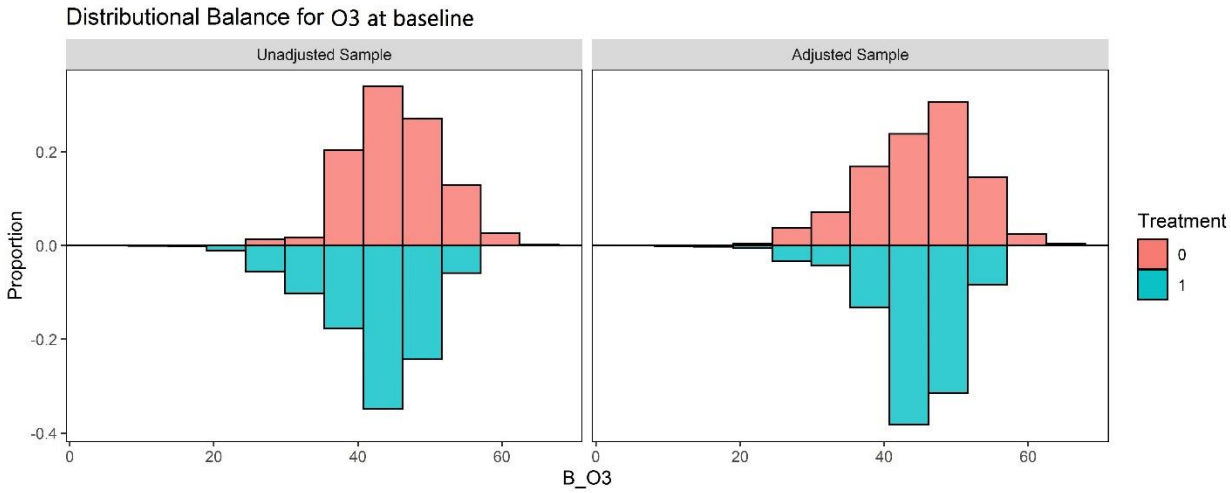

Distributional Balance for education attainment
Unadjusted Sample
Adjusted Sample
Proportion
0.3
0.2
0.1
0.0
1 2 3 4
factor(edu)
Treatment
0
1

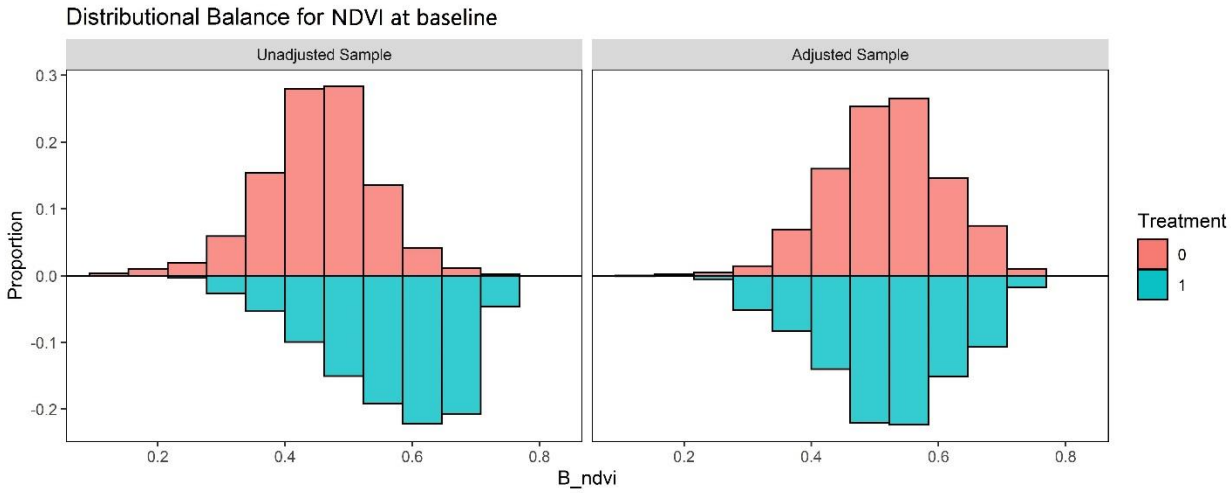

Distributional Balance for occupational class
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
6
factor(occ)
Treatment
0
1

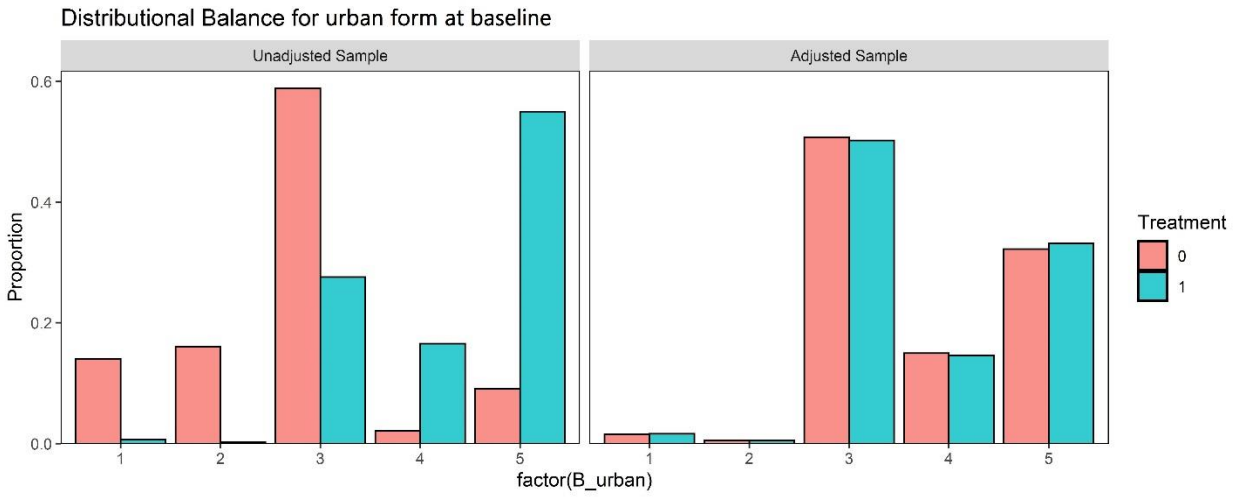

Distributional Balance for annual income (decile) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.05
0.10
0.15
1
2
3
4
5
6
7
8
9
10
factor(B_incdec)
Treatment
0
1

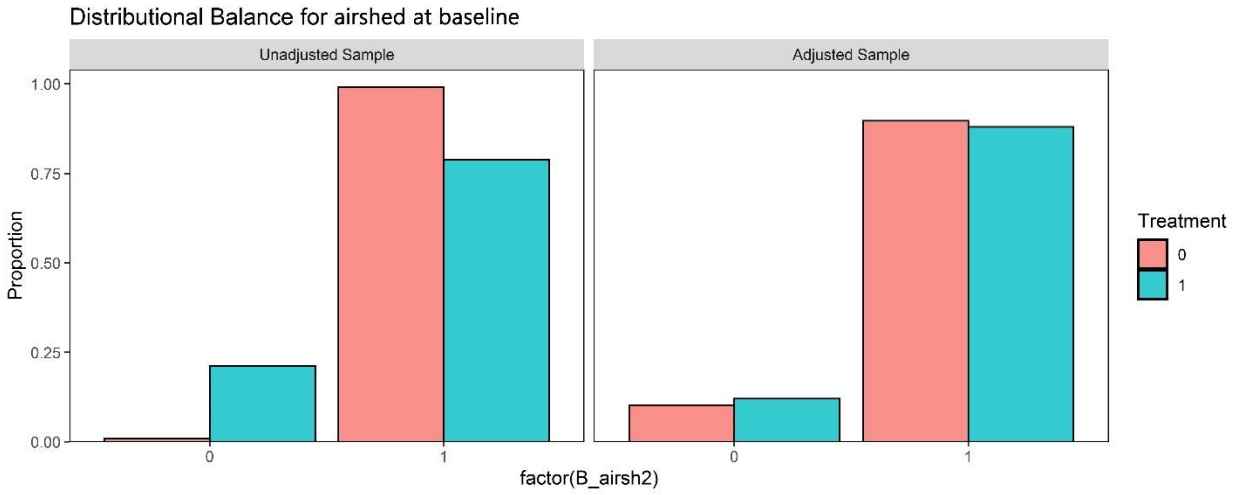

Distributional Balance for Charlson score at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
2
3
factor(B_index_3yr2)
Treatment
0
1

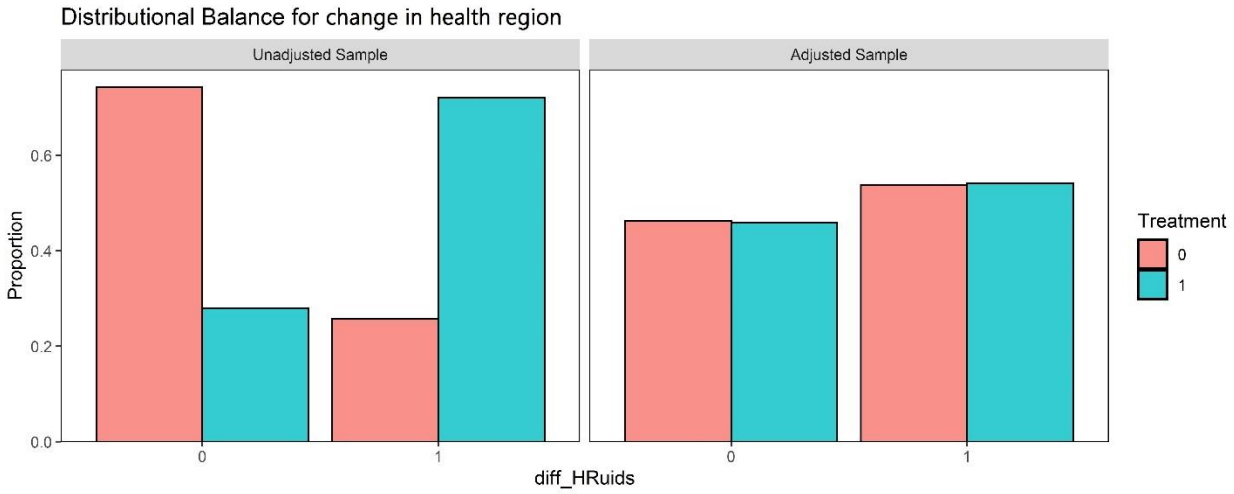

Distributional Balance for NO2 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
0
10
20
30
B_NO2
Treatment
0
1

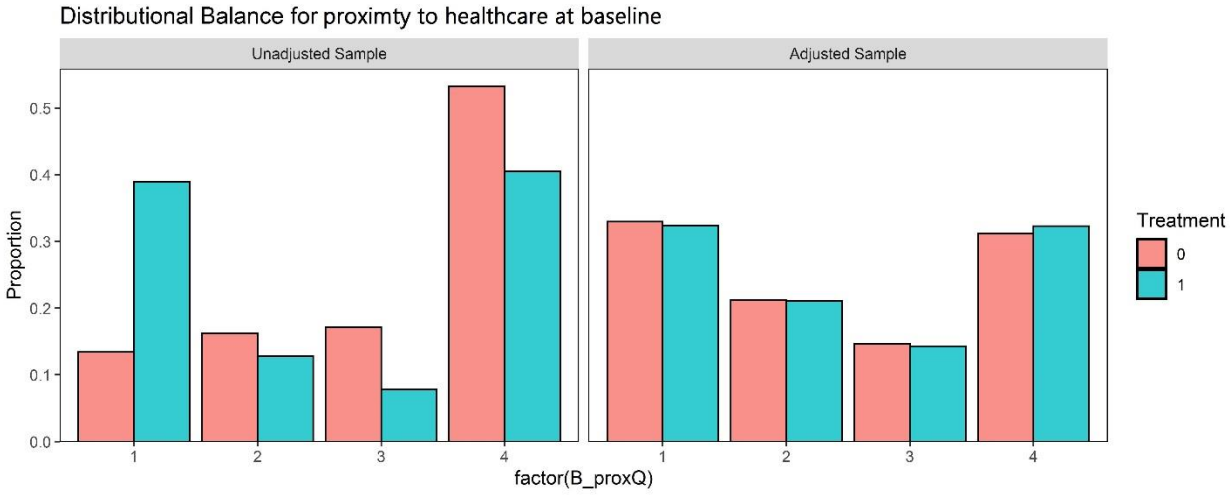

Distributional Balance for O3 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
-0.4
0
20
40
60
B_O3
Treatment
0
1

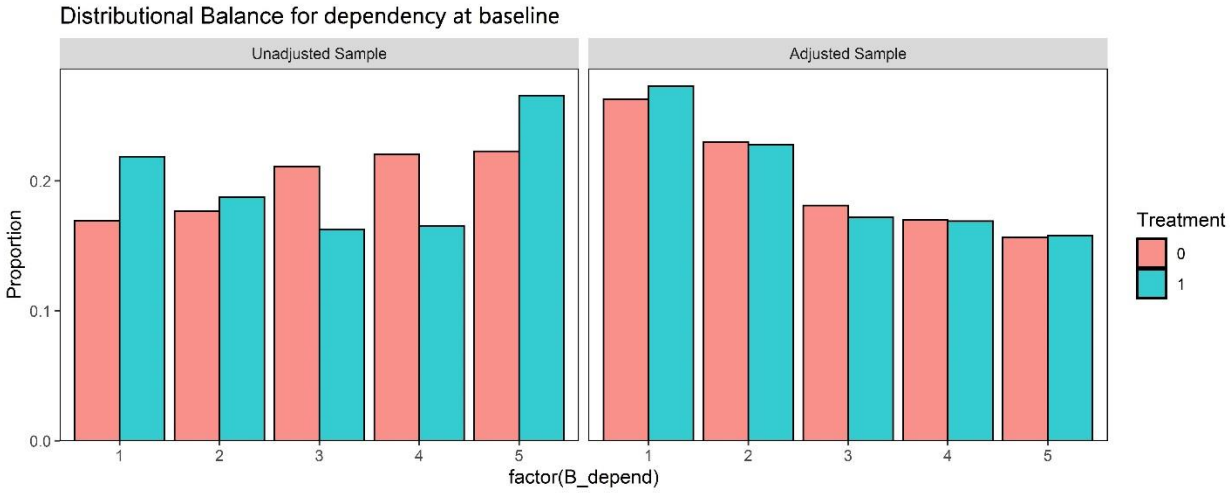

Distributional Balance for NDVI at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.3
0.2
0.1
0.0
-0.1
-0.2
0.2
0.4
0.6
0.8
B_ndvi
Treatment
0
1

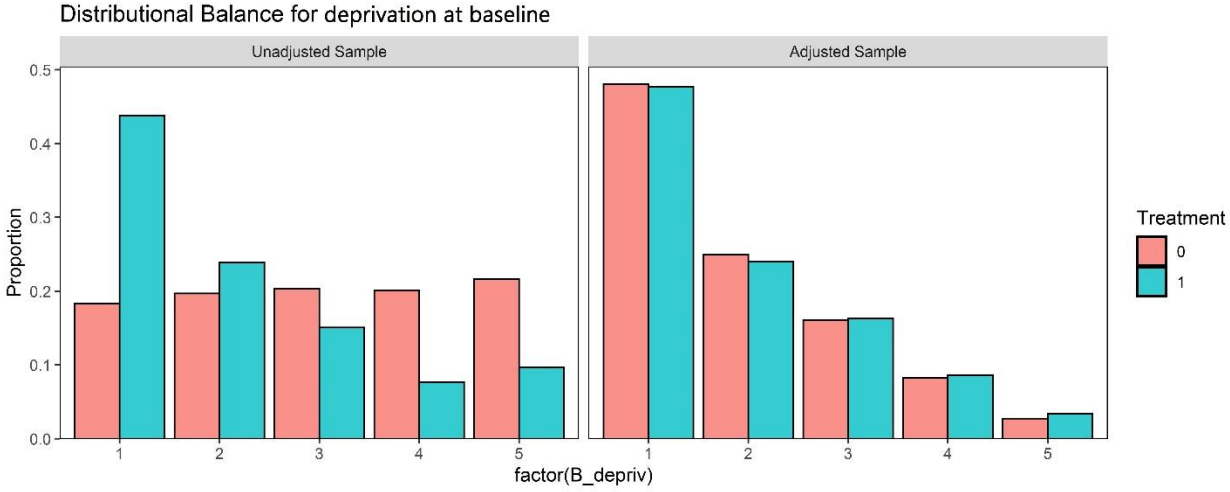

Distributional Balance for urban form at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
1
2
3
4
5
factor(B_urban)
Treatment
0
1

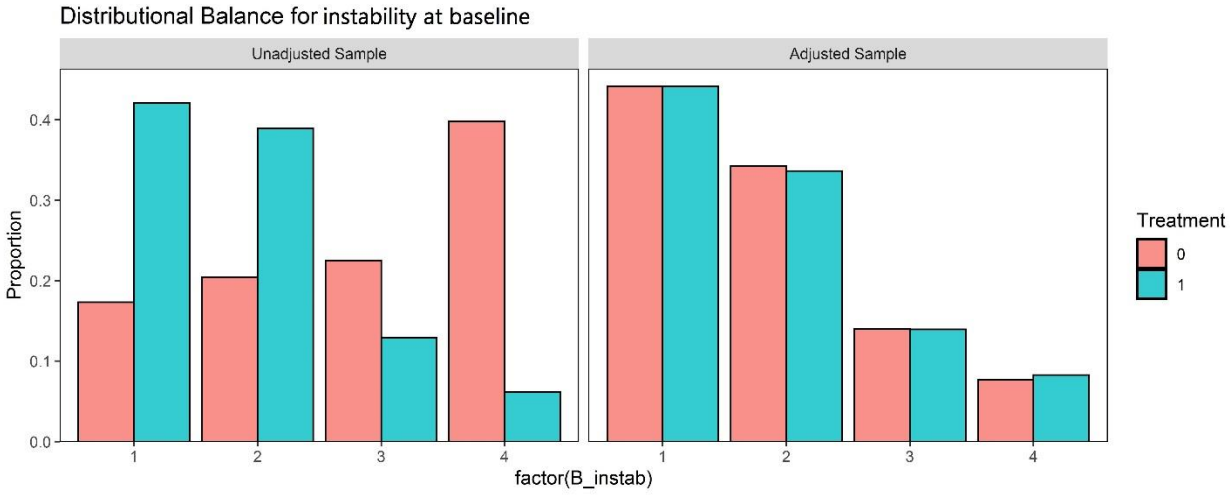

Distributional Balance for airshed at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
1.00
0
1
factor(B_airsh2)
Treatment
0
1

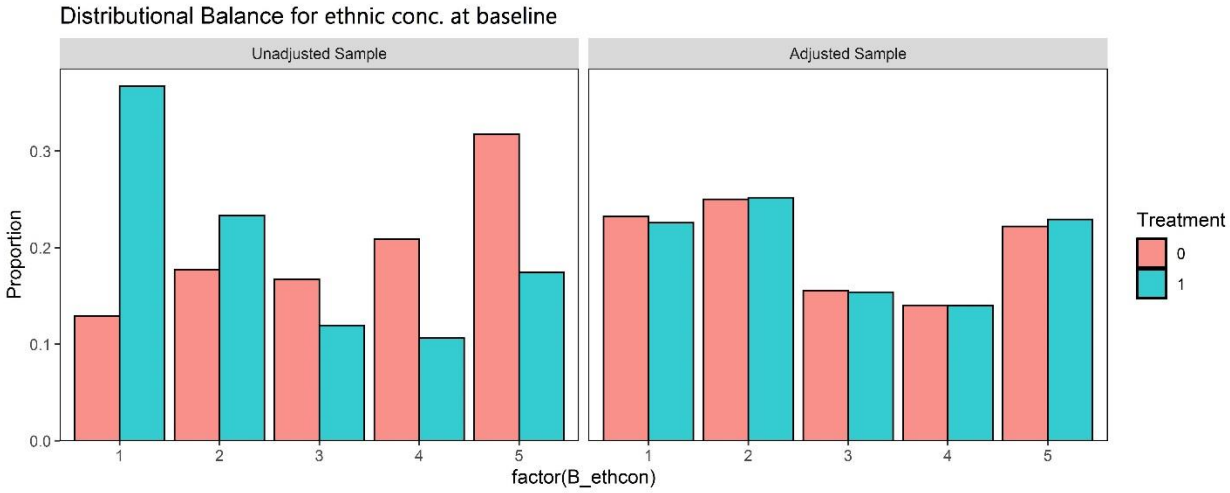

Distributional Balance for change in health region
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0
1
diff_HRuids
Treatment
0
1

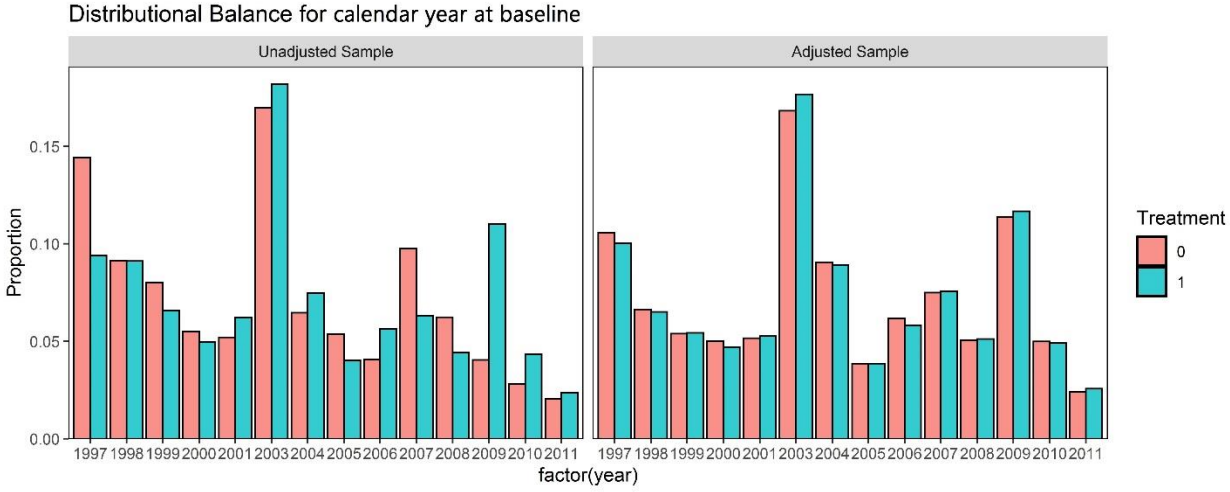

Distributional Balance for proximty to healthcare at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
factor(B_proxQ)
Treatment
0
1


Distributional Balance for dependency at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
factor(B_depend)
Treatment
0
1


Distributional Balance for deprivation at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
factor(B_depriv)
Treatment
0
1


Distributional Balance for instability at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
factor(B_instab)
Treatment
0
1


Distributional Balance for ethnic conc. at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
1
2
3
4
5
factor(B_ethcon)
Treatment
0
1


Distributional Balance for calendar year at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.05
0.10
0.15
1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
factor(year)
Treatment
0
1

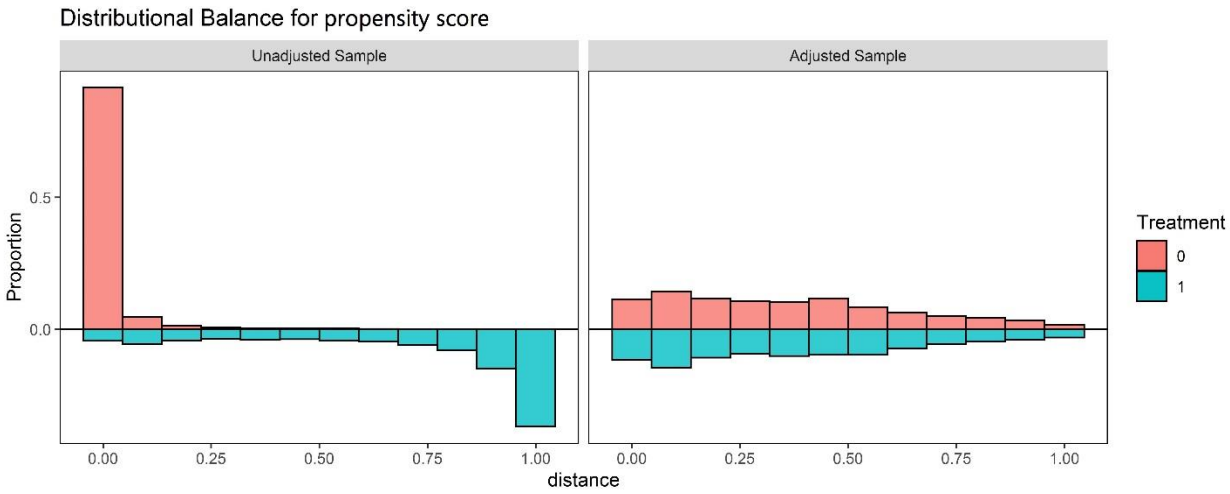
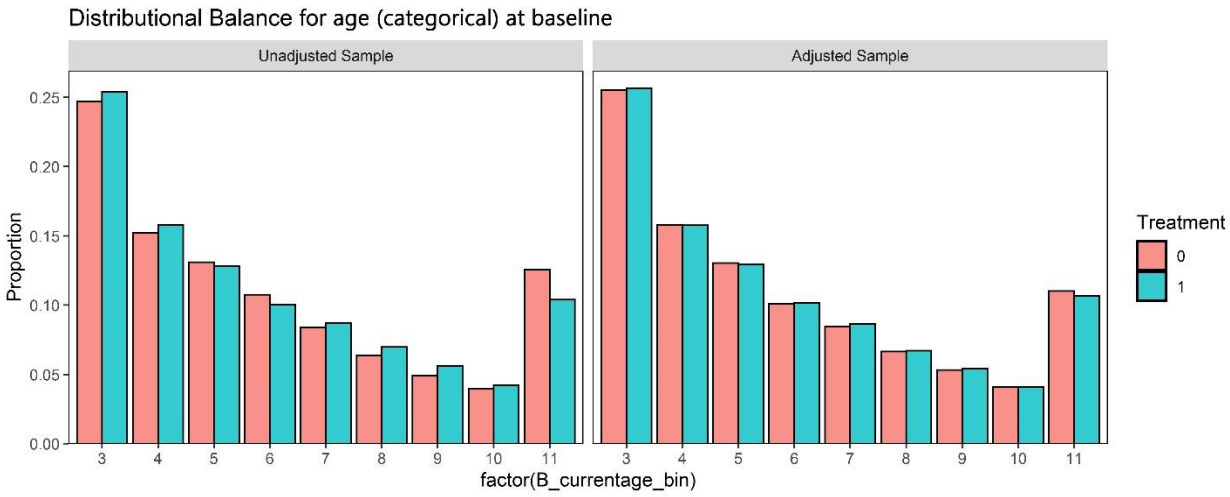
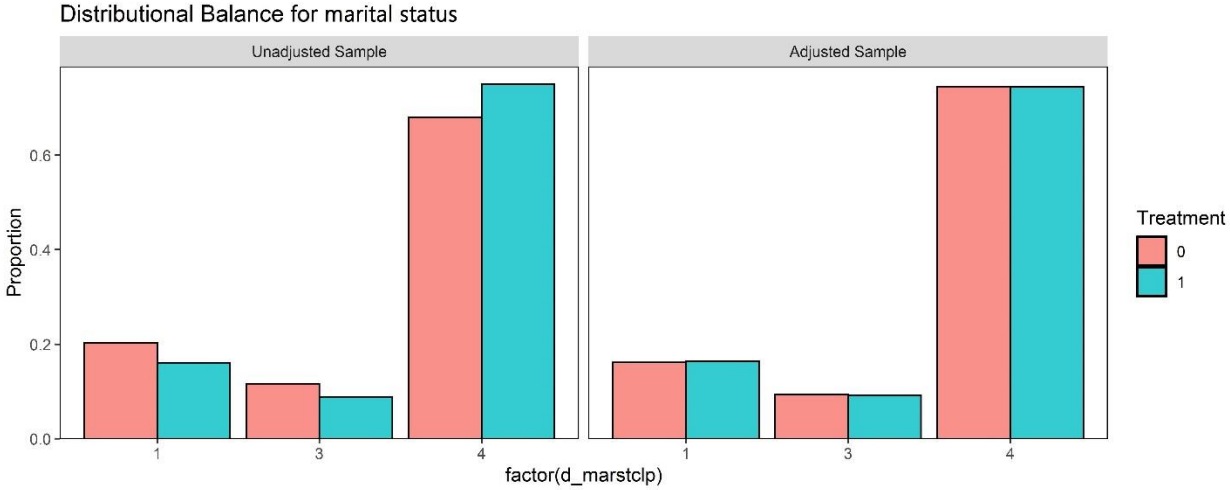
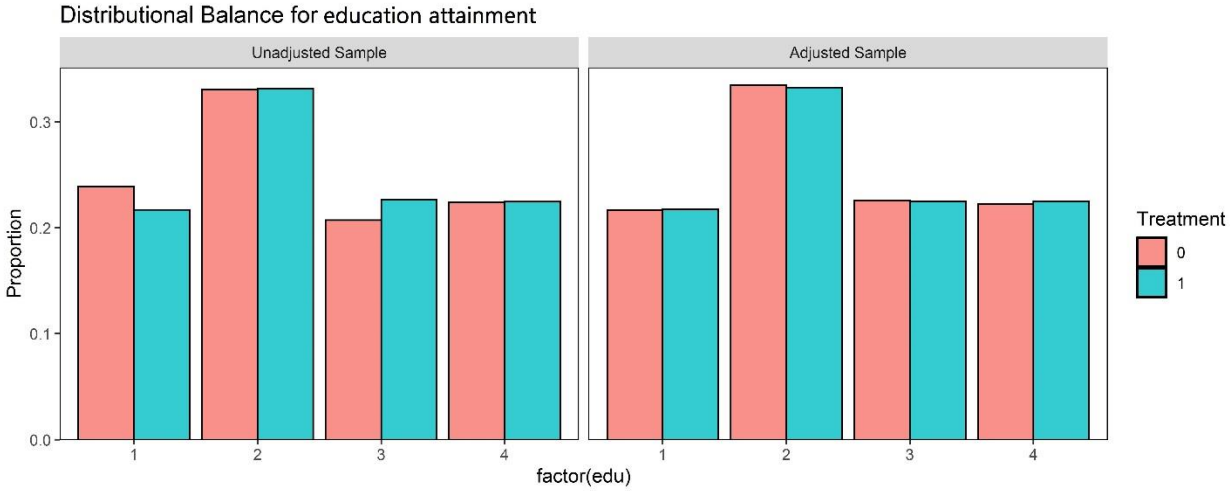
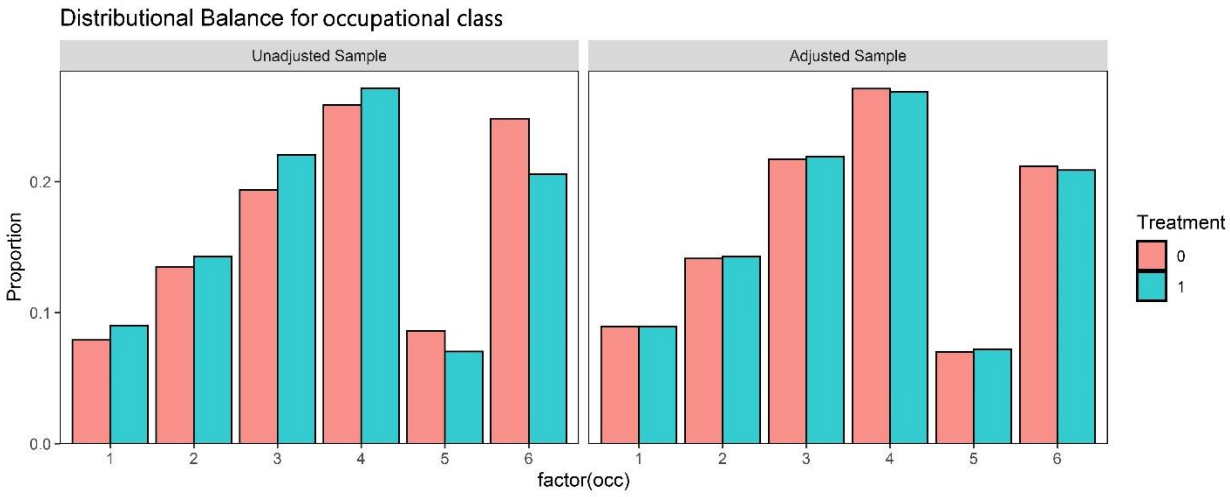


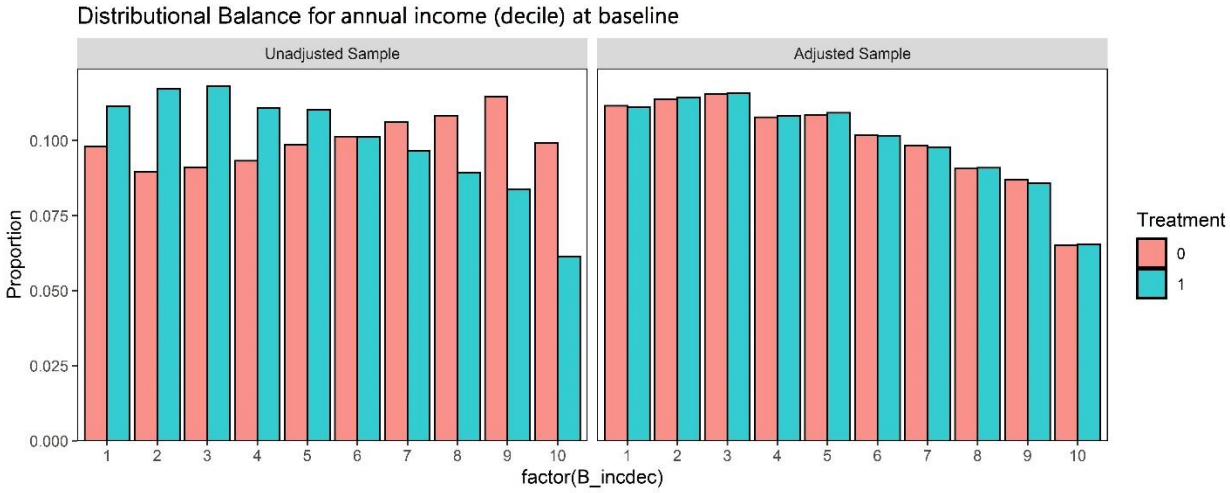

**Figure S4.** Covariate balance between movers from high to low $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

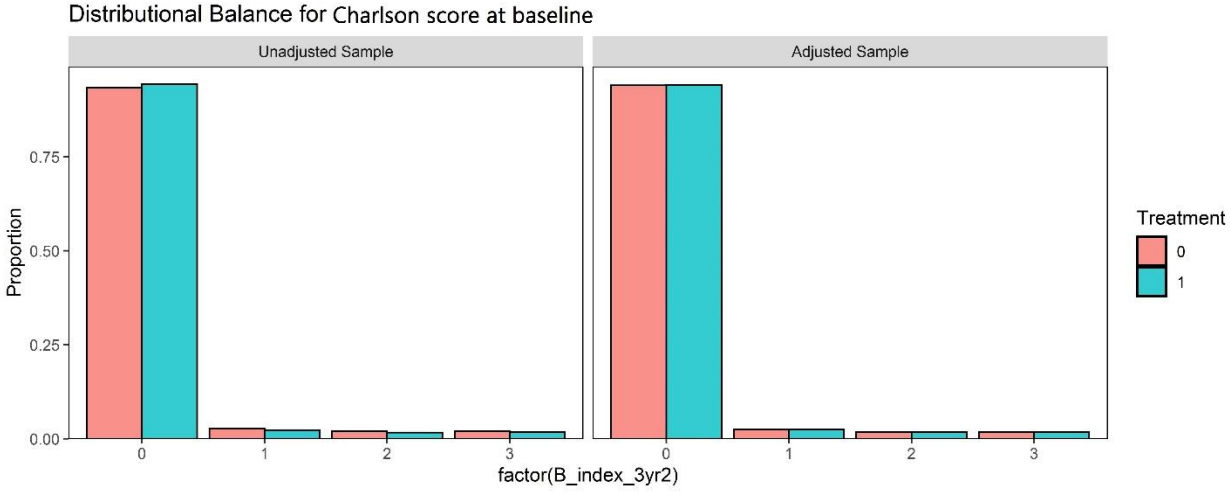

Distributional Balance for age (categorical) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.25
0.20
0.15
0.10
0.05
0.00
3 4 5 6 7 8 9 10 11
3 4 5 6 7 8 9 10 11
factor(B_currentage_bin)
Treatment
0
1

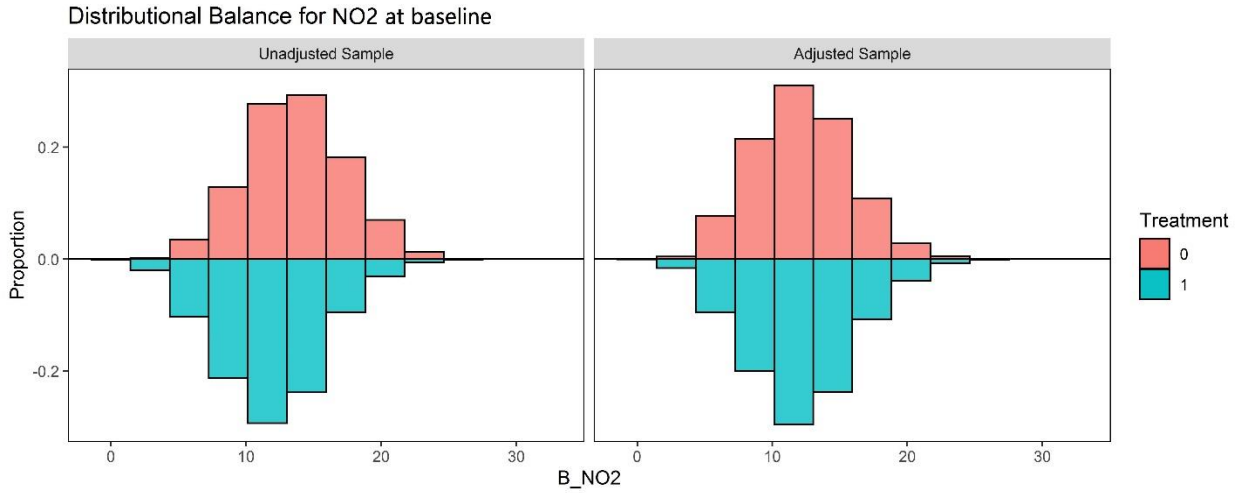

Distributional Balance for marital status
Unadjusted Sample
Adjusted Sample
Proportion
0.6
0.4
0.2
0.0
1 3 4
1 3 4
factor(d_marstclp)
Treatment
0
1

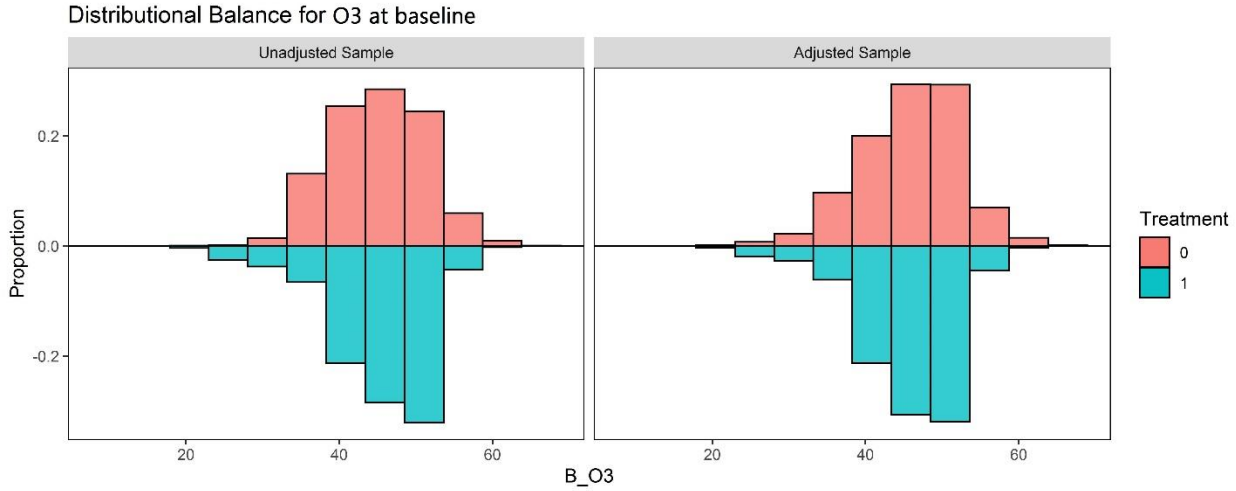

Distributional Balance for education attainment
Unadjusted Sample
Adjusted Sample
Proportion
0.3
0.2
0.1
0.0
1 2 3 4
1 2 3 4
factor(edu)
Treatment
0
1

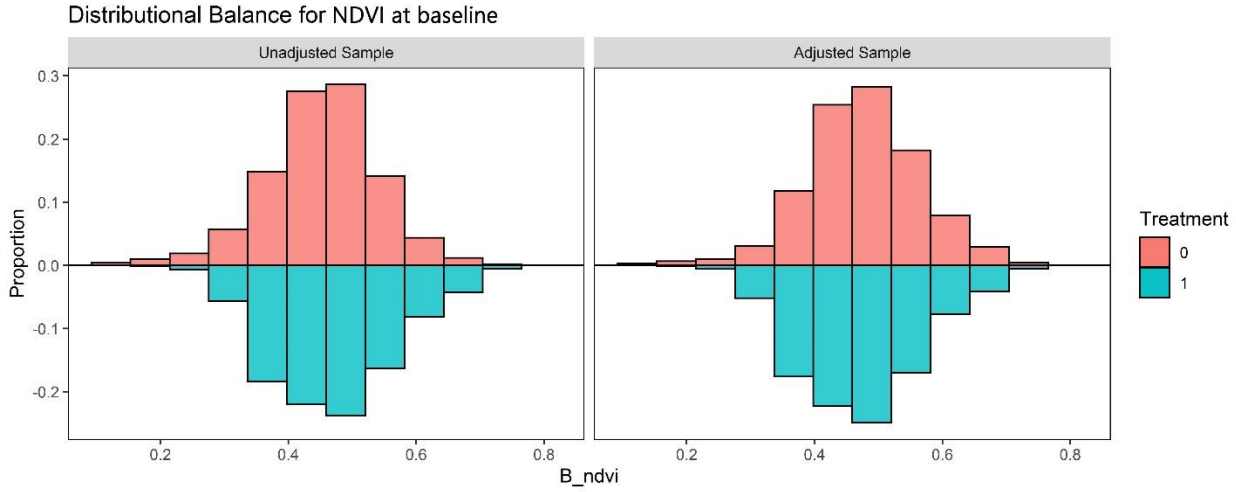

Distributional Balance for occupational class
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
1
2
3
4
5
6
factor(occ)
Treatment
0
1

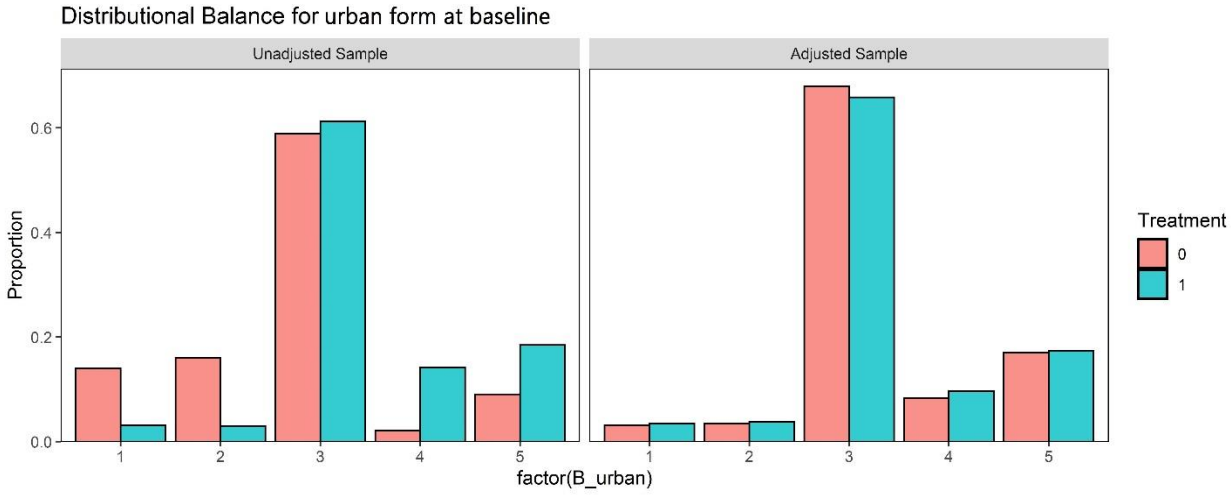

Distributional Balance for annual income (decile) at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.000
0.025
0.050
0.075
0.100
1
2
3
4
5
6
7
8
9
10
factor(B_incdec)
Treatment
0
1

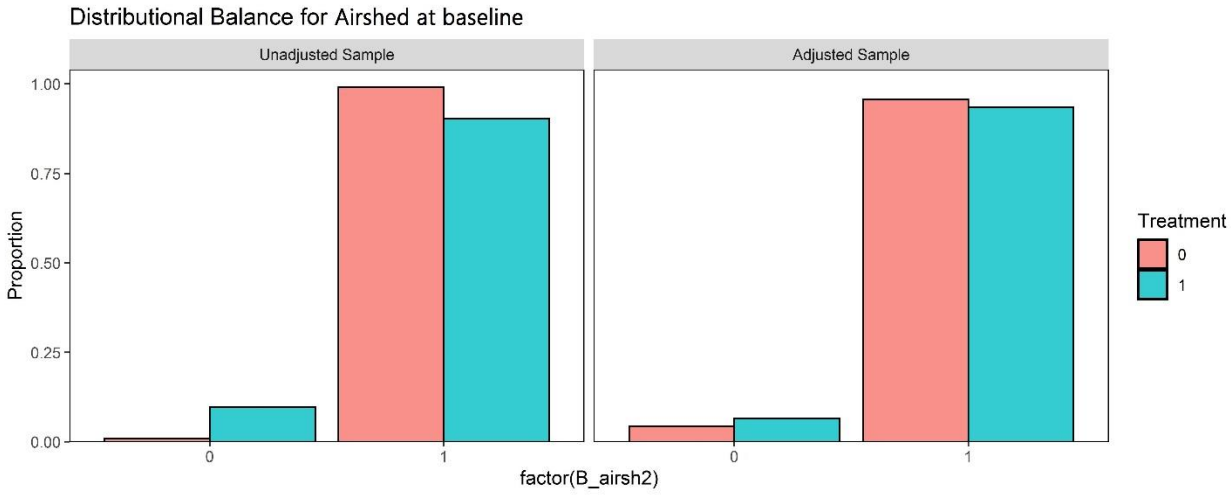

Distributional Balance for Charlson score at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
0
1
2
3
factor(B_index_3yr2)
Treatment
0
1

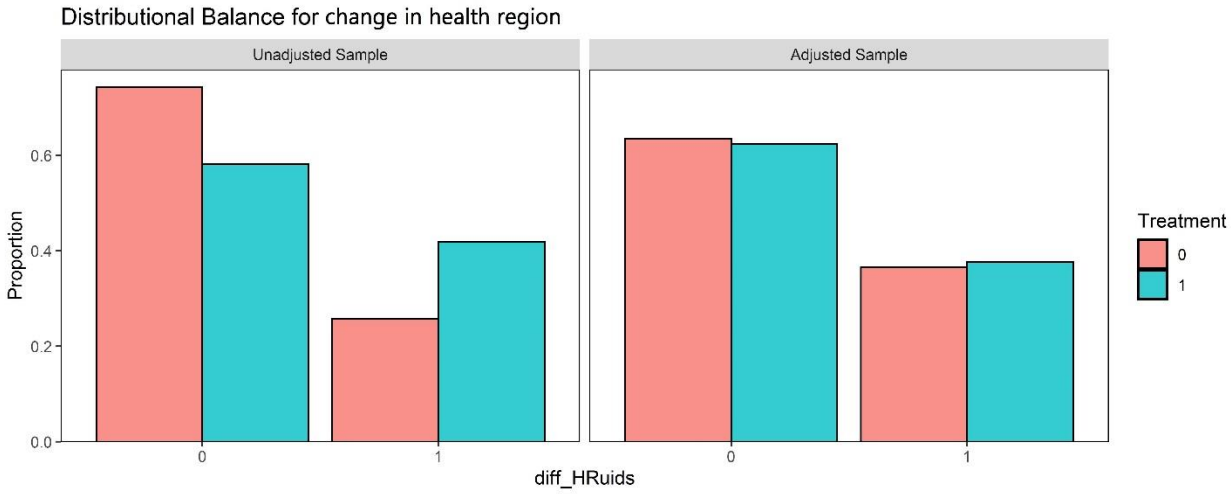

Distributional Balance for NO2 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
0
10
20
30
B_NO2
Treatment
0
1

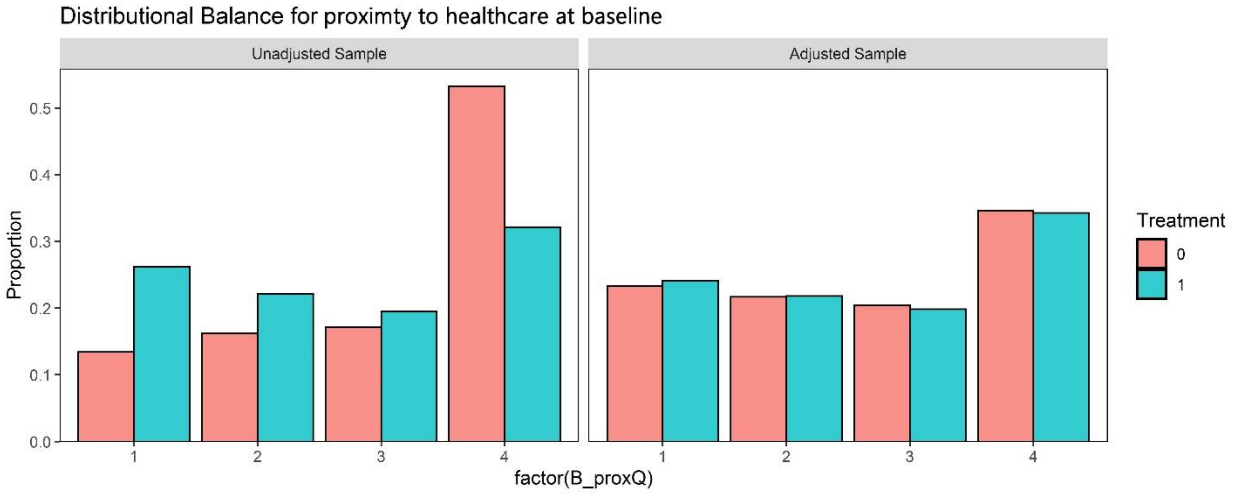

Distributional Balance for O3 at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.2
0.0
-0.2
20
40
60
B_O3
Treatment
0
1

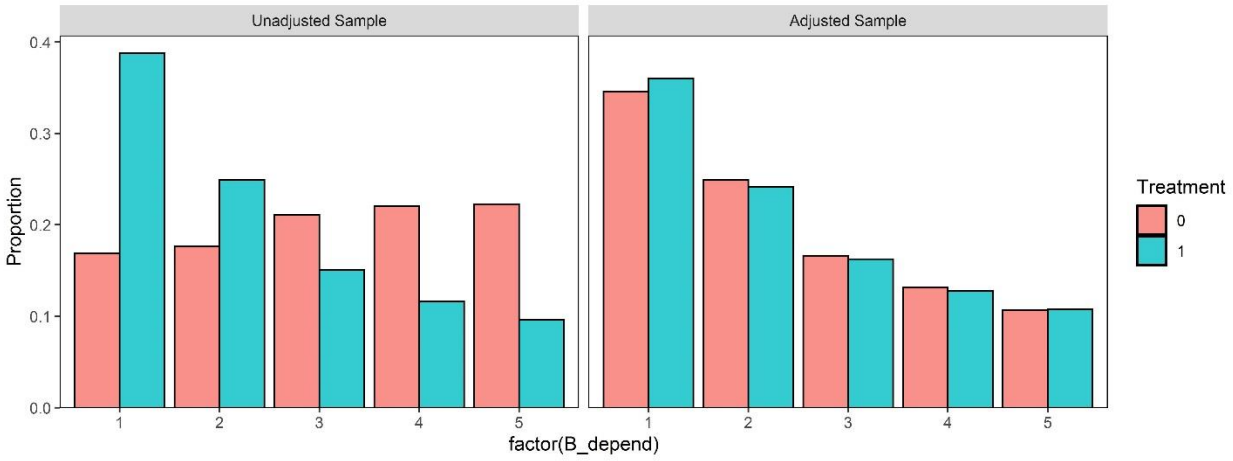

Distributional Balance for NDVI at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.3
0.2
0.1
0.0
-0.1
-0.2
0.2
0.4
0.6
0.8
B_ndvi
Treatment
0
1

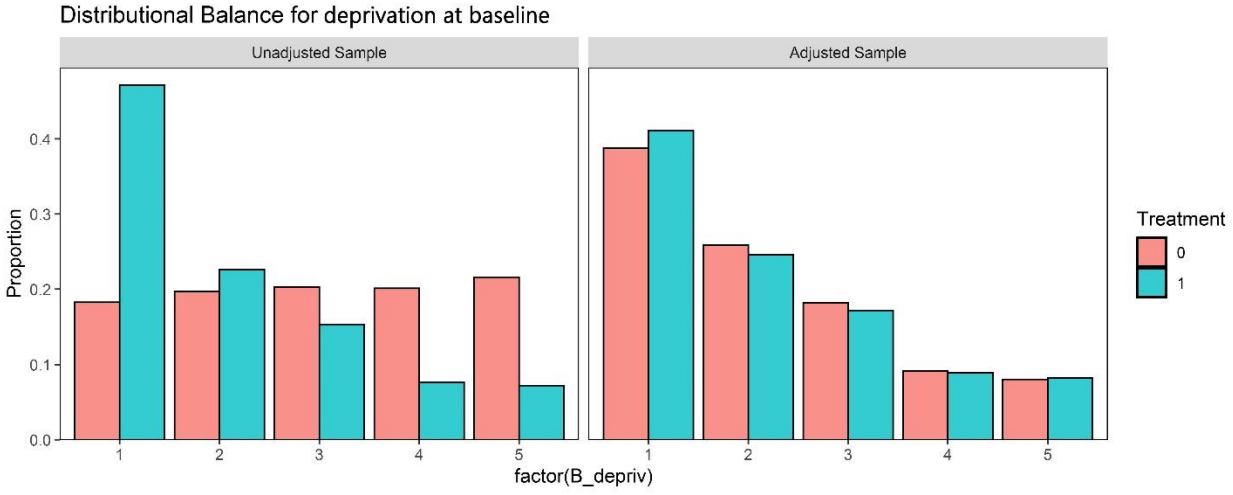

Distributional Balance for urban form at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
1
2
3
4
5
factor(B_urban)
Treatment
0
1

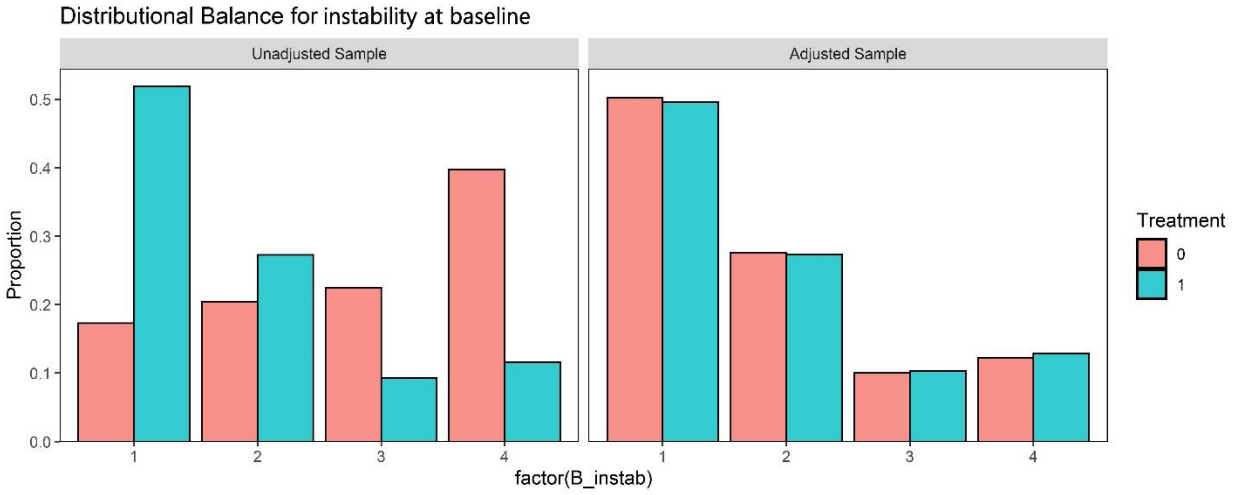

Distributional Balance for Airshed at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.25
0.50
0.75
1.00
0
1
factor(B_airsh2)
Treatment
0
1

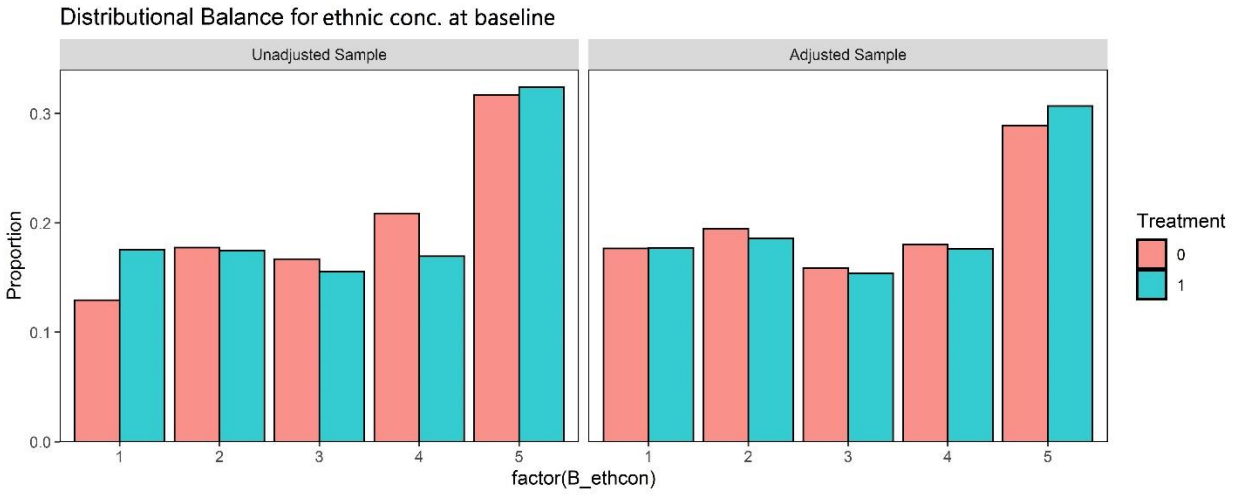

Distributional Balance for change in health region
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.2
0.4
0.6
0
1
diff_HRuids
Treatment
0
1

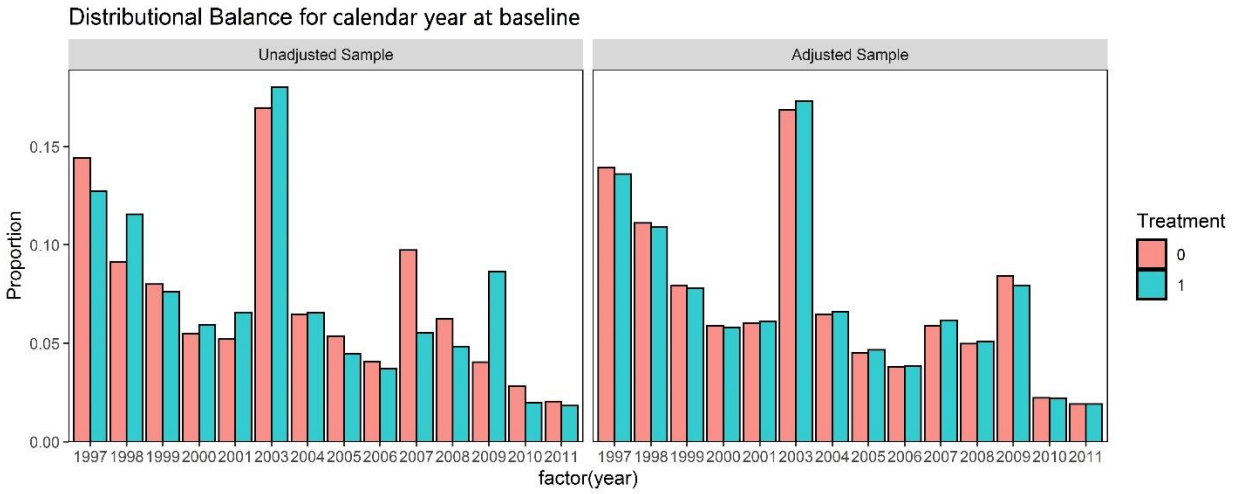

Distributional Balance for proximty to healthcare at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
factor(B_proxQ)
Treatment
0
1


Distributional Balance for dependency at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_depend)
Treatment
0
1


Distributional Balance for deprivation at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
1
2
3
4
5
factor(B_depriv)
Treatment
0
1


Distributional Balance for instability at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
factor(B_instab)
Treatment
0
1


Distributional Balance for ethnic conc. at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.0
0.1
0.2
0.3
1
2
3
4
5
factor(B_ethcon)
Treatment
0
1


Distributional Balance for calendar year at baseline
Unadjusted Sample
Adjusted Sample
Proportion
0.00
0.05
0.10
0.15
1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
factor(year)
Treatment
0
1

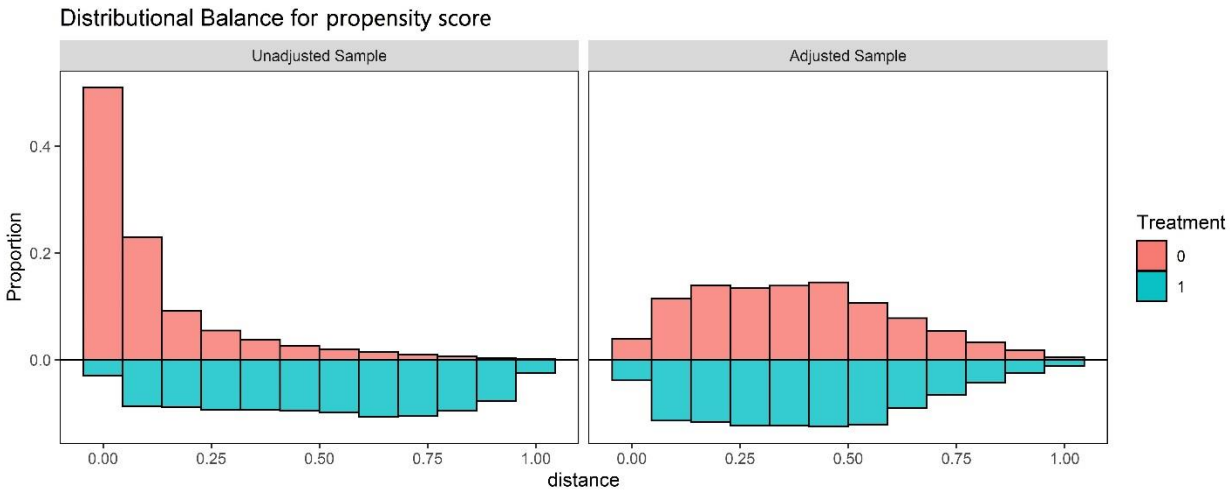



**Figure S5.** Covariate balance between movers from high to intermediate $PM_{2.5}$ areas and their matched controls, before and after applying propensity score matching

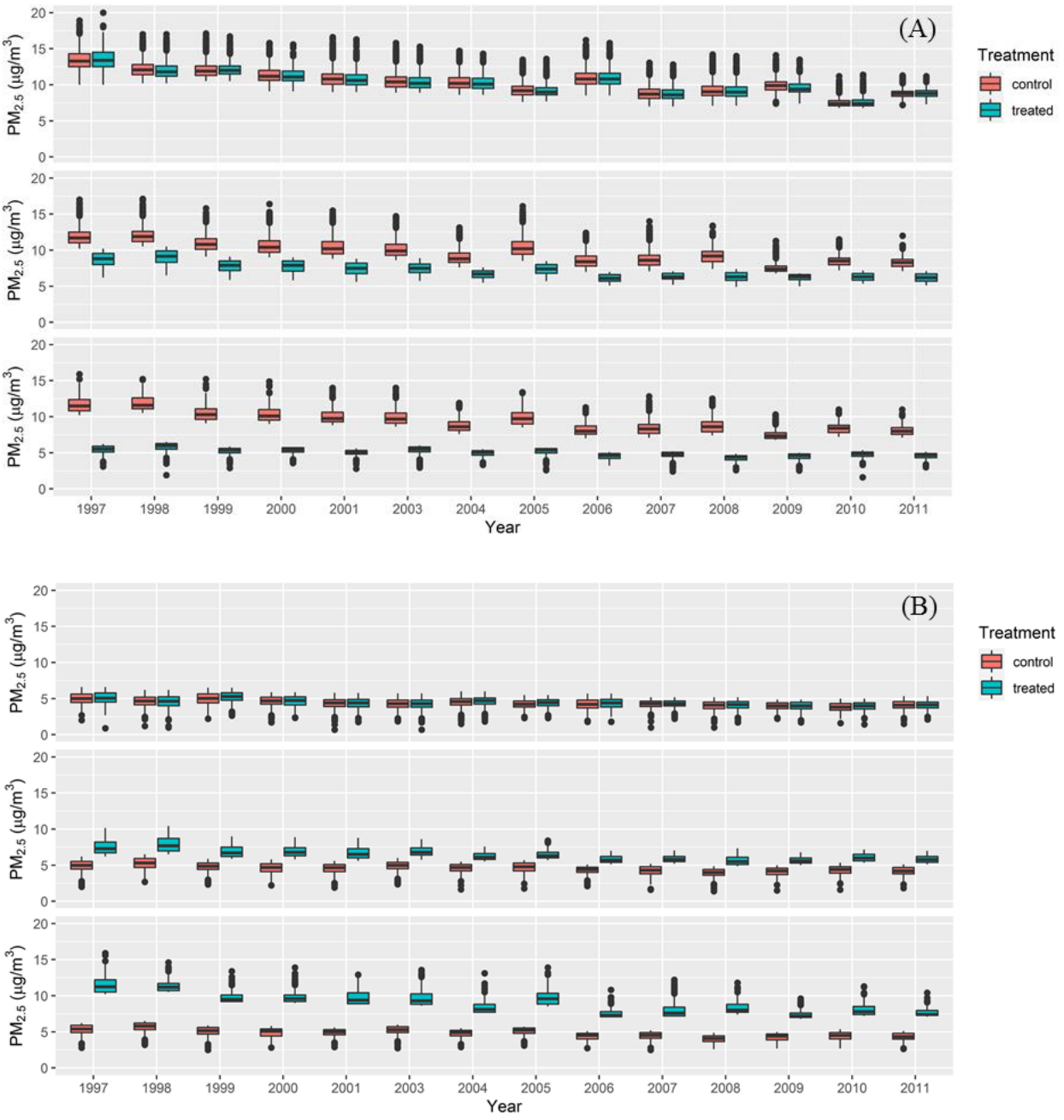



**Figure S6.** Distributions of residential exposure to $PM_{2.5}$ in movers who (A) moved from high to intermediate or low $PM_{2.5}$ areas (referred to as the treated) and their matched controls, as well as in movers who (B) moved from low to intermediate or high $PM_{2.5}$ areas (the treated) and their matched controls, in Canada, 1997 to 2016. For each of the two panels, the upper plot depicts $PM_{2.5}$ exposure before moving, the middle plot depicts $PM_{2.5}$ exposure after moving to intermediate $PM_{2.5}$ areas, and the lower plot depicts $PM_{2.5}$ exposure after moving from either high to low $PM_{2.5}$ or from low to high $PM_{2.5}$ areas

## Reference List


1. Tjepkema M, Christidis T, Bushnik T, Pinault L. Cohort profile: The Canadian Census Health and Environment Cohorts (CanCHECs). *Health reports* 2019; **30**(12): 18-26.
2. Pinault LL, Weichenthal S, Crouse DL, et al. Associations between fine particulate matter and mortality in the 2001 Canadian Census Health and Environment Cohort. *Environ Res* 2017; **159**: 406-15.
3. Rotermann M, Sanmartin C, Trudeau R, St-Jean H. Linking 2006 Census and hospital data in Canada. *Health Reports* 2015; **26**(10): 10.
4. Pappin AJ, Christidis T, Pinault LL, et al. Examining the Shape of the Association between Low Levels of Fine Particulate Matter and Mortality across Three Cycles of the Canadian Census Health and Environment Cohort. *Environ Health Perspect* 2019; **127**(10): 107008.
5. Brauer M, Brook JR, Christidis T, et al. Mortality-Air Pollution Associations in Low-Exposure Environments (MAPLE): Phase 1. *Res Rep Health Eff Inst* 2019; (203): 1-87.
6. Crouse DL, Peters PA, Hystad P, et al. Ambient PM2.5, O3, and NO2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Perspect* 2015; **123**(11): 1180-6.
7. Pope CAr, Burnett RT, Thurston GD, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* 2004; **109**(1): 71-7.
8. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical care* 2005; **43**(11): 1130-9.
9. Roy N, Dubé R, Després C, Freitas A, Légaré F. Choosing between staying at home or moving: A systematic review of factors influencing housing decisions among frail older adults. *PLoS One* 2018; **13**(1): e0189266.
10. Matheson FI, Dunn JR, Smith KL, Moineddin R, Glazier RH. Development of the Canadian Marginalization Index: a new tool for the study of inequality. *Can J Public Health* 2012; **103**(8 Suppl 2): S12-S6.
11. Celis-Morales CA, Lyall DM, Welsh P, et al. Association between active commuting and incident cardiovascular disease, cancer, and mortality: prospective cohort study. *Bmj* 2017; **357**: j1456.
12. Fong KC, Hart JE, James P. A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Current environmental health reports* 2018; **5**(1): 77-87.
13. Markevych I, Schoierer J, Hartig T, et al. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environ Res* 2017; **158**: 301-17.
14. Crouse DL, Pinault L, Balram A, et al. Urban greenness and mortality in Canada's largest cities: a national cohort study. *The Lancet Planetary health* 2017; **1**(7): e289-e97.
15. Chen H, Burnett RT, Bai L, et al. Residential Greenness and Cardiovascular Disease Incidence, Readmission, and Mortality. *Environ Health Perspect* 2020; **128**(8): 87005.
16. Canada S. Proximity Measures Database. 2020. https://www150.statcan.gc.ca/n1/pub/17-26-0002/172600022020001-eng.htm (accessed Oct 1, 2020 2020).
17. Hystad P, Setton E, Cervantes A, et al. Creating national air pollution models for population exposure assessment in Canada. *Environ Health Perspect* 2011; **119**(8): 1123-9.

18. Robichaud A, Menard R. Multi-year objective analyses of warm season ground-level ozone and $PM_{2.5}$ over North America using real-time observations and Canadian operational air quality models. *Atmospheric Chemistry and Physics Discussions* 2013; **13**(5): 13967-4035.
19. Dehejia RH, Wahba S. Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics* 2002; **84**(1): 151-61.
20. Austin PC, Small DS. The use of bootstrapping when using propensity-score matching without replacement: a simulation study. *Stat Med* 2014; **33**(24): 4306-19.
21. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics* 2011; **10**(2): 150-61.
22. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009; **28**(25): 3083-107.
23. Hackam DG, Mamdani M, Li P, Redelmeier DA. Statins and sepsis in patients with cardiovascular disease: a population-based cohort analysis. *Lancet* 2006; **367**(9508): 413-8.
24. Helenius K, Longford N, Lehtonen L, Modi N, Gale C. Association of early postnatal transfer and birth outside a tertiary hospital with mortality and severe brain injury in extremely preterm infants: observational cohort study with propensity score matching. *Bmj* 2019; **367**: l5678.
25. Ouldali N, Toubiana J, Antona D, et al. Association of Intravenous Immunoglobulins Plus Methylprednisolone vs Immunoglobulins Alone With Course of Fever in Multisystem Inflammatory Syndrome in Children. *Jama* 2021; **325**(9): 855-64.
26. Nguyen TL, Collins GS, Spence J, et al. Double-adjustment in propensity score matching analysis: choosing a threshold for considering residual imbalance. *BMC medical research methodology* 2017; **17**(1): 78.
27. Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. *Stat Med* 2013; **32**(16): 2837-49.
28. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *Bmj* 2011; **342**: d40.
29. Montori VM, Guyatt GH. Intention-to-treat principle. *Cmaj* 2001; **165**(10): 1339-41.
30. Bacon SL, Lavoie KL, Bourbeau J, et al. The effects of a multisite aerobic exercise intervention on asthma morbidity in sedentary adults with asthma: the Ex-asthma study randomised controlled trial protocol. *BMJ Open* 2013; **3**(6).
31. Eftekhari A, Ruzek JI, Crowley JJ, Rosen CS, Greenbaum MA, Karlin BE. Effectiveness of National Implementation of Prolonged Exposure Therapy in Veterans Affairs Care. *JAMA Psychiatry* 2013.
32. Hiyoshi A, Fukuda Y, Shipley MJ, Brunner EJ. Inequalities in self-rated health in Japan 1986-2007 according to household income and a novel occupational classification: national sampling survey series. *J Epidemiol Community Health* 2013.
33. Ton TG, Longstreth WT, Jr., Koepsell T. Active and passive smoking and risk of narcolepsy in people with HLA DQB1*0602: a population-based case-control study. *Neuroepidemiology* 2009; **32**(2): 114-21.
34. Hindorff LA, Rice KM, Lange LA, et al. Common variants in the CRP gene in relation to longevity and cause-specific mortality in older adults: the Cardiovascular Health Study. *Atherosclerosis* 2008; **197**(2): 922-30.

35. Clark TG, Altman DG. Developing a prognostic model in the presence of missing data: an ovarian cancer case study. *J Clin Epidemiol* 2003; **56**(1): 28-37.
36. Chase AJ, Fretz EB, Warburton WP, et al. Association of the arterial access site at angioplasty with transfusion and mortality: the M.O.R.T.A.L study (Mortality benefit Of Reduced Transfusion after percutaneous coronary intervention via the Arm or Leg). *Heart* 2008; **94**(8): 1019-25.
37. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. *Circulation* 2005; **112**(2): 224-31.
38. Schafer JL. Analysis of incomplete multivariate data. London: Chapman & Hall; 1997.
39. Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.
40. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods* 2001; **6**(4): 330-51.
41. Meng XL. Multiple-imputation inferences with uncongenial sources of input. *Statistical Science* 1994: 538-58.
42. van BS, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999; **18**(6): 681-94.
43. Little RJA. Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics* 1988; **6**(3): 287-96.
44. Marshall A, Altman DG, Holder RL. Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study. *BMC Med Res Methodol* 2010; **10**: 112.
45. Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of statistical software* 2010: 1-68.
46. Van Buuren S, Brand JPL, Groothuis-Oudshoorn CGM, Rubin DB. Fully conditional specification in multivariate imputation. *Journal of statistical computation and simulation* 2006; **76**(12): 1049-64.
47. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med* 2009; **28**(15): 1982-98.
48. van BS, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations. R package version 3.13.0. *http://CRAN R-project org/package=mice*, 2021, 2021. https://cran.r-project.org/web/packages/mice/mice.pdf (accessed.