Appendix 1

We consider and describe further the causal relationships for the example discussed in the text, a spatial study of the health effects of ambient air pollution, that is using *area-specific* air pollution levels (e.g., ambient PM 2.5), as represented by AP_0 in the directed acyclic graph[1] shown in Figure 1A. The health outcome is the individual birth weight (D_1 in Figure 1A) of infants born during the study period in these areas. The study goal is to assess the effects, if any, of air pollution levels on birth weight represented by the dashed arrow. Measured covariates (e.g., C_0 in Figure 1A) affect birth weight. If these covariates are also associated with air pollution levels (AP_0) as shown in Figure 1A, they will tend to confound effect estimates and need analytic control. This tendency to confound results is represented in Figure 1A by the open path from exposure to disease through C_0.

Unmeasured or misclassified factors present prior to birth are represented by U_0 in Figure 2A. For example, U_0 might be unmeasured or inadequately measured poverty. If U_0 affects both birth weight and exposure, say because poor people might tend to live nearer to pollution sources, then U_0 would tend to confound estimates. This tendency is reflected by the open path from exposure to disease through U_0. The path between exposure and disease through C_0 is blocked, if we control for C_0 analytically (indicated by a box around C_0) in Figure 2A. Even if we suspected from previous knowledge that U_o could confound the $AP_0 – D_1$ relationship but had been *un*able to measure it – we would not be able to support empirically that confounding was present. Our goal is to show how consideration of *future* pollution levels may allow empiric detection of residual confounding even if the responsible confounder is unmeasured or unrecognized.

To show how we can detect residual confounding, we include in Figure 3A future air pollution (AP_2) which occurs in a *later*, second time period, strictly after the first period during which the outcome (D_1) occurs and is measured. No arrow goes from AP_2 to the health outcome, since it cannot be a cause (due to the temporal sequence); further, no arrow goes from D_1 to AP_2 since we *assume* that the outcome does not affect air pollution. We have boxed in AP_0 and C_0 to indicate analytic control for them, say by

stratification. The arrow from U_0 to AP_0 and AP_2 indicates that (we assume) U_0 affects both (e.g., pollution sources continue to be nearer to poor neighborhoods, even in the *later* time period). Any confounding causes of AP_2 also cause AP_0. With the indicated causal relationships, we would expect to find an association between AP_2 and the outcome, i.e. a path through U_0 if confounding is present (an arrow from U_0 to D_1) but no association if confounding were absent (no arrow from U_0 to D_1 or no arrows from U_0 to AP_0 and AP_2). The same arguments hold if we also include an arrow from AP_0 to AP_2 to reflect the possibility that earlier exposures cause later ones (Figure 4A). If a factor, such as U_0 in Figure 5A affects disease (D_1) and future air pollution (AP_2), but is not associated with the exposure of interest (AP_0), the outcome could be associated with the indicator even in the absence of residual confounding. Thus, we must carefully consider the causal relationships and evaluate the likelihood that a factor, perhaps like U_0 in Figure 5A, could create an association between the future indicator and disease even in the absence of residual confounding.

If the causal relationships are like those in Figures 1A-4A, these arguments suggest that we can use a factor, such as air pollutant levels for a period after health events have already occurred, as an indicator of unmeasured confounding. Briefly stated, our central assertion is: If unmeasured confounding is present and if our basic causal assumptions reasonably approximate reality (e.g., Figures 1A-4a), then future air pollution (AP_2) may be associated with past outcomes if there is residual confounding, whereas in the absence of unmeasured confounding or other model mis-specification, AP_2 should be independent of past outcomes conditional on AP_0.

Our arguments emphasize residual confounding – our focus. This can be viewed as model misspecification where important causal factors are omitted leading to confounding. Although the proposed method is intended to identify residual confounding, other types of model misspecification (e.g. if the indicator correlates with an important factor measured with error) can also result in an association between the "indicator" variable and outcome, and the indicator cannot distinguish between them[2,3].

Figure 1A

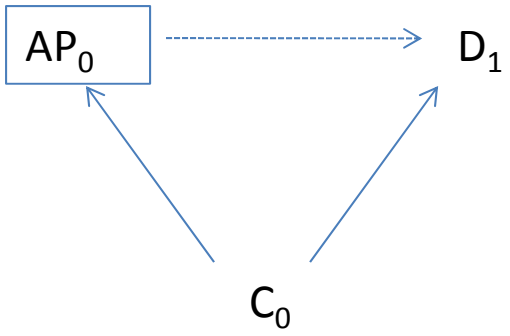

Figure 2A

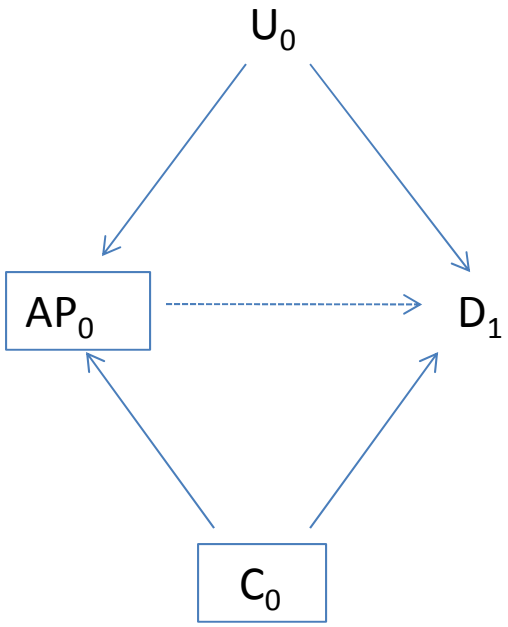

Figure 3A

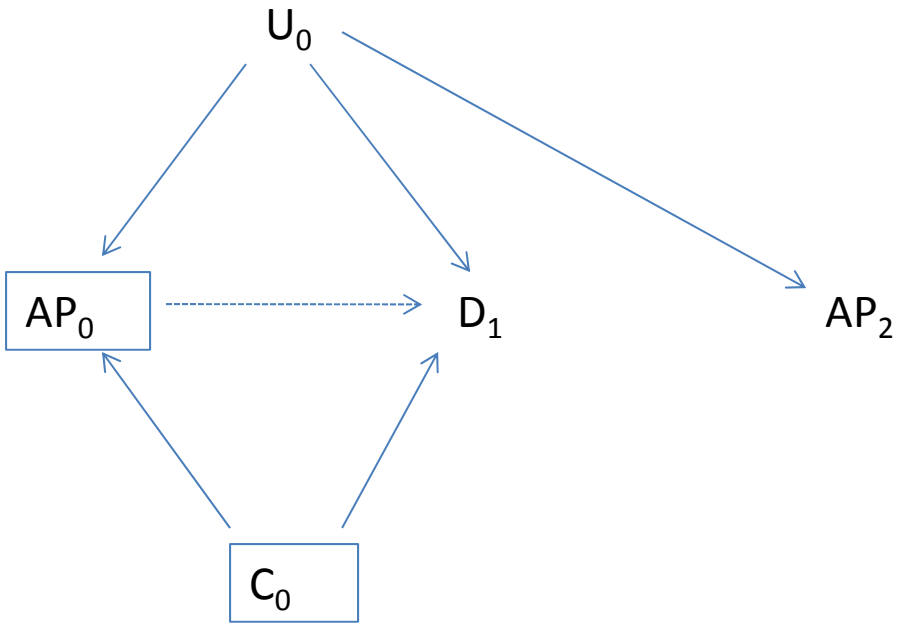

Figure 4A

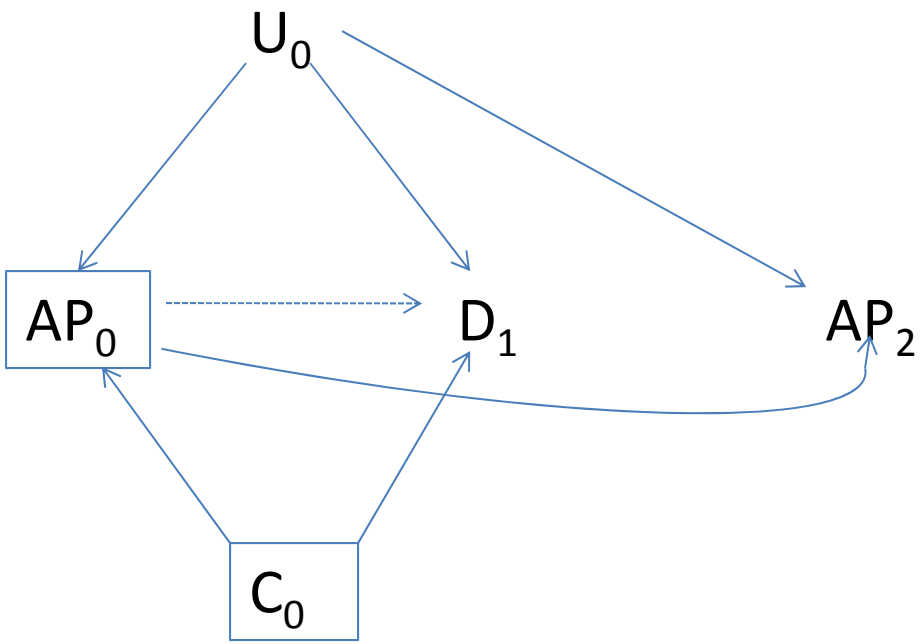

Figure 5A