Beyond the Diagonal Reference Model: Critiques & New Directions in the Analysis of Mobility Effects

Ethan Fosse
University of Toronto

and

Fabian T. Pfeffer
Ludwig-Maximilians-University Munich

Keywords: Social mobility, Diagonal reference model,
Social stratification, Mobility effects, Social class

Corresponding Author:

Ethan Fosse
Assistant Professor of Sociology
Associate Director, Data Sciences Institute
University of Toronto
ethan.fosse@utoronto.ca

Abstract

Over the past decade there has been a striking increase in the number of quantitative studies examining the effects of social (i.e., socioeconomic) mobility, with almost all recent results in sociology and demography based on Sobel's (1981, 1985) Diagonal Reference Model (DRM). This paper makes four main contributions to this rapidly expanding literature. First, we show that under plausible values of mobility effects, the DRM will, in general, implicitly force the underlying mobility linear effect to zero. In addition, we show both mathematically and through simulations that the mobility effects estimated by the DRM are sensitive to the size and sign of the origin and destination linear effects, often in ways that are unlikely to be intuitive to applied researchers. This finding clarifies why, contrary to expectations, applied researchers have generally found weak or no evidence of mobility effects on a wide range of outcomes. Second, we generalize the identification problem of conventional mobility effect models by showing that the DRM and related methods can be viewed as special cases of a bounding analysis, where identification is achieved by invoking extremely strong assumptions (resulting in very tight bounds). Finally, and importantly, we present a new framework for the analysis of mobility tables based on the identification and estimation of joint parameter sets, introducing what we call the Structural and Dynamic Inequality (SDI) model. We show that this model is fully identified, relies on much weaker assumptions than conventional models of mobility effects, and can be treated both as a descriptive model and, if additional assumptions are invoked, as a causal model. We conclude with an agenda for further research on the consequences of socioeconomic mobility.

Introduction

The inconsistent (and often small) mobility effects that this literature has estimated appear to refute longstanding sociological theory and present a general challenge to much of the work on social mobility undertaken by sociologists and demographers. Since the influential work of Lipset and Bendix (1959), researchers have carried out a number of important and widely cited studies exploring differences in rates of social mobility across and within countries (e.g., Bloome 2015; Chetty et al. 2014; Erikson and Goldthorpe 1992; Grusky and Hauser 1984). To the extent that a high degree of social fluidity is considered a normatively desirable aim (for a critical perspective see Swift 2004), these studies provide important descriptive evidence on the degree of openness of a given society or place. However, as Lipset and Zetterberg (1959) pointed out, "unless variations in mobility rates and in the subjective experience of mobility make a difference for society or for the behavior pattern of an individual, knowledge concerning rates of mobility will be of purely academic interest" (6). Based on today's quantitative evidence, it would seem that generations of sociologists and demographers have been trying to explain an outcome that is, on the whole, of little consequence to the individual. Interest in the consequences of social mobility is longstanding (for reviews see Hendrickx et al. 1993; Hope 1971, 1975). From its earliest days, the scientific literature on social mobility, or the experience of moving up or down the class hierarchy of a given society, has debated its individual-level effects. For example, the sociologist Pitirim Sorokin (1927) hypothesized negative effects of not just downward but also upward social mobility on individual well-being, as those who reach a status different from that of their parents may suffer from the cultural gap between their attained position and their family origins (see also Friedman 2016). Nearly a century later, there is a large body of quantitative sociological research that has sought to estimate the direct effects of the experience of social mobility on a wide range of individual outcomes, including well-being (e.g., life satisfaction, stress, allostatic load, substance abuse), attitudes (e.g., trust, political ideology, redistribution preferences), and behaviors (e.g., voting, fertility, health behaviors) (see also Online Appendix D).

Progress in empirically identifying the effects of social mobility on individuals has been hampered by a fundamental methodological challenge. Observed social mobility (M) is simply the difference between an individual's social destination (D), such as their own social class, and their social origin (O), such as their parents' social class, so that $M = D - O$. As a result, any model of mobility effects that seeks to estimate the independent effects of social origins, social destinations, and the difference between them, i.e., social mobility, is underidentified and cannot be estimated using conventional statistical techniques (Manski 1990, 2003). In contrast to problems of statistical inference, which involve understanding how sampling variability can affect conclusions based on samples of limited size, problems of identification entail understanding what conclusions can be drawn even with a sample of unlimited size. The lack of a unique solution of mobility effects is a classic identification problem, because it cannot be resolved by collecting larger samples.

To estimate such effects, a variety of techniques have been proposed, but the most popular approaches are those developed by the sociologists Otis Dudley Duncan (1966) and Michael Sobel (1981, 1985). A first wave of research, influenced by Duncan's (1966) *Square Additive Model* (SAM), a basic two-factor origin-destination model with residual interaction terms, found no effect of social mobility on a number of outcomes (see also Hope 1971, 1975). This is partly explained by the fact that Duncan's proposed model assumes that the linear effect of mobility is zero, as was recognized by at least some social scientists at the time (Blalock 1967: 794-795). Another wave of research resulted from Sobel's (1981, 1985) *Diagonal Reference Model* (DRM),[1], which is seen as the "gold standard" for mobility effects research (Houle and Martin 2011: 197; Präg and Richards 2018:5; Sieben 2017). The statistician Sir David Cox (1990), for example, has lauded the DRM as an exemplar of "directly substantive" models in statistical analysis (170). In recent years, sociologists have used Sobel's model to explain a wide range of outcomes, including subjective well-being, political extremism, obesity, and so on (the list is too long to include here but see Online Appendix D). However, as we show, like Duncan's model, the DRM relies on very strong assumptions about the effects of social mobility.

The remainder of this paper is organized as follows. First, we outline the identification challenge, clarifying what can be known about the data from a mobility table with as few assumptions as possible. We make explicit how, under a general model of mobility effects, the nonlinear effects are identified and the linear effects are not. Second, we discuss the mathematical properties of the DRM, revealing that the DRM generates different mobility linear effects depending on both the size and sign of the linear and nonlinear effects of origin and destination. In doing so, we revisit Sobel's (1981) findings on fertility and show that his data are consistent with a wide range of very large negative as well as positive linear mobility effects, none of which are recovered by the DRM. Third, we generalize models of mobility effects, showing how any existing mobility model that attempts to identify unique "effects" can be viewed as a special case of a bounding approach, except with extremely narrow bounds and thus extremely strong assumptions. Finally, we outline a new framework for analyzing mobility data using what we call the *Structural and Dynamic Inequality* (SDI) model. Using this model, we demonstrate how one can describe both dynamic (or mobility-based) inequalities as well as those that are purely structural (reflecting an absence of social mobility). As we discuss, these estimates can also be interpreted as joint causal effects, and therefore involve much weaker assumptions than estimates from conventional models of mobility effects. We conclude with a programmatic statement outlining guidelines for further research on the individual-level consequences of social mobility.

[1] In this paper we use "diagonal reference model" to refer to Sobel's "simple diagonal reference model" (1981), which includes a single weight for the entire mobility table. However, our main claims hold for other versions of the DRM that allow weights to differ by either origin or destination class (Sobel 1985), or by both origin and destination (Weakliem 1992: 157).

I. The Diagonal Reference Model

As noted in the introduction, by far the most common approach to model mobility effects is the *Diagonal Reference Model* (DRM) developed by the sociologist Michael Sobel (1981, 1985). This popularity is indicated by the spike in recent works using the DRM to estimate mobility effects, as shown in Figure 1. This is also demonstrated by the widespread agreement in the literature on the importance and utility of the DRM for understanding the consequences of social mobility. For instance, Houle and Martin (2011: 197) note, the DRM is "the only method used in modern mobility effects research." Likewise, Sieben (2017) writes that DRMs "are thought to be the best solution to [the identification] problem." Präg and Richards (2018: 5) echo this claim, correctly stating: "A consensus is emerging in the literature that the diagonal reference model is superior to other modelling approaches and results based on other approaches are questionable at best."

There is similarly widespread agreement that the DRM is effective in separating out the effects of mobility from those of origin and destination. For example, Schuck and Steiber (2018: 1249) write that the DRM "tests for the net effects of intergenerational mobility *over and above* the effects of educational origin and destination, finally allowing mobility effects to be separated from mere level effects [emphasis in original]." Unfortunately, however, the DRM relies on extremely strong assumptions and, absent additional information external to the data, there is no reason to think that the model will actually recover the "true" underlying mobility effect.

In this section, we first outline the basic mathematical properties of the DRM. As we illustrate, the DRM assumes that the origin and destination effects are proportional to each other, and thus we will refer to this assumption as the "proportionality constraint" (see also Weakliem 1992: 157). Failure to satisfy this constraint may lead to erroneous estimates of the overall effects of origin, destination, and mobility. This is true of the linear effects, but also of the nonlinear effects, which, as noted above, are identified.[2] Second, we revisit Sobel's (1981) fertility data and show, mathematically and with simulations, that the DRM will fix the mobility linear effect at a very specific value even when the underlying mobility linear effect is extremely positive or negative. This is another way of stating that the proportionality constraint may hold for the nonlinear effects, but still not

[2]This means that this constraint can be partially tested against the data. For a similar point, see Weakliem (1992: 157).

hold for the linear effects. Unfortunately, the data are not themselves informative about whether or not this constraint is satisfied. Finally, we develop a simple formula that, when the proportionality constraint is satisfied, clarifies the nature of the bias of the estimated mobility linear effect. As we show, the mobility linear effect generated by DRM is a function of the size, sign, and shape of the origin and destination effects. It is only under very specific conditions, namely when the ratio of the origin and destination weights equals the ratio of the underlying origin and destination linear effects, that the DRM will recover the true underlying mobility linear effect.

1. Overview of the DRM

To orient the discussion that follows, suppose we have a set of categorical variables for $i = 1, \ldots I$ origin groups, $j = 1, \ldots, J$ destination groups, and $k = j - i + I, \ldots K$ mobility groups. The fundamental model of mobility effects can be specified using what we call the *Classical Origin-Destination-Mobility (C-ODM) model*:

$$Y = f(O^*, D^*, M^*) + \epsilon = \mu + \alpha_i + \beta_j + \gamma_k + \eta_{ijk} + \xi_{rijk}, \tag{1}$$

where μ is the intercept (or overall mean); α_i, β_j, γ_k denote the ith, jth, kth observed levels of origin, destination, and mobility, respectively; η_{ijk} is an additional (orthogonal) term denoting interactions; and ξ_{rijk} is an individual-level, normally-distributed error term with a mean of zero. As we discuss in Online Appendix B, Equation 1 is based on the implicit assumption that O, D, and M (and their respective indices) can be treated as surrogates for distinct causal variables O^*, D^*, and M^* (and their respective indices). To simplify the exposition, we will accordingly refer to α_i, β_j, and γ_k as the "true" origin, destination, and mobility effects, but the reader should keep in mind that this is shorthand, as detailed in Online Appendix B, for referring to the causal effects α^*_l, β^*_p, γ^*_n (for a similar point, see Fosse and Winship 2019a).

An alternative formulation of the C-ODM model (see Equation 1) helps clarify the nature of the identification problem. By orthogonalizing the linear from the nonlinear terms, we can specify what we call the *Linearized Origin-Destination-Mobility (L-ODM) model*:

$$\mu_{rijk} = \mu + \alpha(i - i^*) + \beta(j - j^*) + \gamma(k - k^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j + \widetilde{\gamma}_k + \eta_{ijk} + \xi_{rijk} \tag{2}$$

where the asterisks denote midpoint or referent indices $i^* = (I+1)/2$, $j^* = (J+1)/2$, and $k^* = (K+1)/2$; α, β, and γ denote the linear effects of origin, destination, and mobility, respectively; and $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$ represent the origin, destination, and mobility nonlinear effects, respectively; η_{ijk} is, as before, an additional (orthogonal) term denoting interactions; and ξ_{rijk} is a normally-distributed individual-level error term with a mean of zero. To identify the levels of the parameters given the inclusion of the intercept, sum-to-zero constraints are applied to the linear and nonlinear parameters. The L-ODM greatly simplifies the nature of the identification problem, allowing us to show how the DRM makes highly specific assumptions about the unknown origin, destination, and mobility linear effects. Moreover, as we show later, it allows one to use graphical tools for visualizing and partially identifying the parameters of a mobility effects model.

The L-ODM highlights that the core identification challenge resides in the linear effects (α, β, γ). The Diagonal Reference Model (DRM), the most prevalent method used to address this challenge and estimate distinct mobility effects, achieves point identification by imposing highly specific identifying constraints with respect to these linear effects. To understand this, it is crucial to recognize that the DRM is a nonlinear model in that at least one of the (unknown) model parameters are a nonlinear function of at least one of the other (unknown) model parameters. By contrast, conventional linear regression is linear in the parameters, although it can be parameterized so that it is nonlinear in the variables. The general form for a nonlinear regression model is as follows:

$$y_r = f(\mathbf{x}_r, \boldsymbol{\theta}) + \epsilon_r \tag{3}$$

where $r = 1, \ldots, R$ indexes the rows (e.g., individuals) of the data set, $q = 1, \ldots, Q$ the number of variables in the model, and $p = 1, \ldots, P$ the number of parameters; y_r is the outcome of the rth row; f is a known function; $\mathbf{x}_r$ is a $Q \times 1$ column vector of variables $x_{1r}, \ldots, x_{Qr}$; $\boldsymbol{\theta}$ is a $P \times 1$ column vector of parameters $\theta_1, \ldots, \theta_P$; and ϵ_r is an error term.[3]

Suppose we have a data set based on aggregated data from a mobility table with rows indexed by $r = 1, \ldots, R = I \times J$. The DRM is a particular kind of nonlinear model that begins with the following as a baseline model:

[3]Similar to conventional linear regression, it is typically assumed that the errors are uncorrelated with each other and have a mean of zero with constant variance.

$$\mu_{ijk} = f(\mathbf{x}_r, \boldsymbol{\theta}) = w_o \mu_{i[j=i]} + w_d \mu_{[i=j]j} + \phi_{ijk} \tag{4}$$

with

$$w_o = \frac{e^{\xi_1}}{e^{\xi_1} + e^{\xi_2}} \text{ and } w_d = 1 - w_o = \frac{e^{\xi_2}}{e^{\xi_1} + e^{\xi_2}},$$

where μ_{ijk} is the mean outcome for a particular cell on the mobility table indexed by i and j (with $k = j - i + I$); w_o is the class origin weight; w_d is the class destination weight, which is equal to $1 - w_o$; ξ_1 and ξ_2 are the origin and destination parameters, respectively, used to calculate the weights; $\mu_{i[j=i]}$ and $\mu_{[i=j]j}$ are the diagonal means of the mobility table indexed by class origin and destination, respectively; and ϕ_{ijk} is a cell-specific error term.[4]

Above the baseline model, mobility effects can be parameterized as a set of categorical variables:

$$\mu_{ijk} = w_o \mu_{i[j=i]} + w_d \mu_{[i=j]j} + \gamma_k + \phi_{ijk} \tag{5}$$

where γ_k denotes the set of mobility effect parameters using, say, sum-to-zero deviation coding. Sobel (1981: 902) suggests that mobility effects are present whenever the γ_k's in Equation 5 are statistically significant.[5]

The DRM as outlined in Equation 5 and applied to a three-by-three mobility table is shown in Table 1. The joint origin and destination effects for the cells on the main diagonal, representing non-mobile groups, are given by the estimates of $\mu_{i[j=i]}$ (or, equivalently, $\mu_{[i=j]j}$).[6] For example, if there are $I = 3$ origin categories and $J = 3$ destination categories, as is the case in Table 1, then the joint origin and destination effects on the main diagonal are given by μ_{11}, μ_{22}, and μ_{33}. The joint

[4]Note that, because Equation 4 is nonlinear, it cannot be estimated using ordinary least squares. Formally, for nonlinear models, at least one of the derivatives of the expectation function with respect to the parameters depends on at least one or more of the parameters.

[5]In the analyses below, we assume that we have included a full set of nonlinear effects for mobility. In practice, many researchers have tended to aggregate the mobility effects variables (see Sobel 1981: 901-902). In fully-identified models this is not particularly problematic, but in the case of unidentified models such seemingly minor decisions can have important consequences. Note that collapsing just two of mobility categories is sufficient to identify the model, equivalent to making a very strong, tacit assumption about the size and sign of the underlying linear effects. Collapsing a larger number of categories has an equivalent impact. These seemingly minor modeling decisions contribute in part to the heterogeneity of results in the mobility effects literature.

[6]To see this, let the main diagonal cells be indexed by $n = 1, \ldots, N$. For nth cell on the main diagonal, where $\mu_{i[j=i]} = \mu_{[i=j]j} = \mu_n$, the joint origin and destination effects equal $w_o\mu_n + w_d\mu_n = w_o\mu_n + (1 - w_o)\mu_n = \mu_n$.

origin and destination effects for the cells off the main diagonal, representing mobile groups, are given by $w_o\mu_{i[j=i]} + w_d\mu_{[i=j]j}$, where w_o and w_d can be interpreted as the relative salience of the origin and destination categories, respectively. For instance, as shown in Table 1, the joint effects for origin and destination in the cell $i = 2$ and $j = 1$ are given by $w_o\mu_{22} + w_d\mu_{11}$.

To understand the underlying mathematical properties of the DRM, it is useful to describe the weights needed to recover the underlying origin, destination, and mobility effects.[7] The true means on the main diagonal can be written as (cf. Equation 2):

$$\mu_{i[j=i]} \text{ (or } \mu_{[i=j]j}) = \mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j \quad \text{with } j = i \text{ (or } i = j), \tag{6}$$

where μ is the overall mean; α and β are the true origin and destination mobility linear effects; $\widetilde{\alpha}_i$ is the ith origin nonlinear effect; $\widetilde{\beta}_i$ is the jth destination nonlinear effect; $i^* = (I + 1)/2$ and $j^* = (J + 1)/2$. Note again that we assume that, to identify the intercept, we have applied sum-to-zero constraints on the origin, destination, and mobility effects.

The weight required to recover the true contribution of the ith origin effect is:

$$w_{o_i} = \frac{\alpha(i - i^*) + \widetilde{\alpha}_i}{\mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j} \qquad \text{with } j = i, \tag{7}$$

where w_{o_i} is the weight of the ith origin category. Then multiplying this weight by the main diagonal mean (Equation 6) will generate the ith origin effect:

$$\mu_{i[j=i]}w_{o_i} = \left(\mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j\right)\left(\frac{\alpha(i - i^*) + \widetilde{\alpha}_i}{\mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j}\right) = \alpha(i - i^*) + \widetilde{\alpha}_i \quad \text{with } j = i.$$

Similarly, the actual weight required to recover the true value of the jth destination effect is:

$$w_{d_j} = \frac{\beta(j - j^*) + \widetilde{\beta}_j}{\alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j} \qquad \text{with } i = j, \tag{8}$$

where w_{d_j} is the weight of the jth destination category. Multiplying this weight by the main diagonal mean (Equation 6) will generate the jth destination effect:

$$\mu_{[i=j]j}w_{d_j} = \left(\mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j\right)\left(\frac{\beta(j - j^*) + \widetilde{\beta}_j}{\mu + \alpha(i - i^*) + \beta(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j}\right) = \beta(j - j^*) + \widetilde{\beta}_j \quad \text{with } i = j.$$

[7] For simplicity, and without a loss of generality, we assume that there are no error terms in the following sections.

Crucially, the weights in Equations 7 and 8 are unknown, because they require knowledge of the unidentified origin and destination linear effects. Note further that the weights are allowed to vary across class and destination origins, respectively, and there is no restriction that the weights sum to one (or some other value).

By contrast, the DRM's proportionality constraint means that we assume that the origin and destination effects can be characterized by a single origin weight (or, equivalently, destination weight). In essence, the DRM attempts to recover the true origin and destination effects (and, accordingly, the true mobility effects) by splitting the main diagonal cells of a mobility table into origin and destination components. More specifically, the proportionality constraint assumes that the relative contributions of the linear and nonlinear origin effects with $i = j$ are the same, and that the same relative proportion of the grand (or overall) mean is attributable to class origin. That is, with respect to the class origin weight, the DRM assumes the following:

$$w_o = \frac{\mu_\alpha}{\mu} = \frac{\alpha}{\alpha + \beta} = \frac{\widetilde{\alpha}_i}{\widetilde{\alpha}_i + \widetilde{\beta}_{[j=i]}} \quad \text{for } i = 1, \ldots, I, \tag{9}$$

where μ_α is that part of the overall mean that is attributable to class origin. Similarly, the DRM assumes that the relative contributions of the linear and nonlinear destination effects with $j = i$ are identical, with the same relative proportion of the grand (or overall) mean attributable to class destination. In other words, the DRM assumes the following with respect to the class destination weights:

$$w_d = \frac{\mu_\beta}{\mu} = \frac{\beta}{\alpha + \beta} = \frac{\widetilde{\beta}_j}{\widetilde{\alpha}_{i=j} + \widetilde{\beta}_j} \quad \text{for } j = 1, \ldots, J, \tag{10}$$

where μ_β is that part of the overall mean that is attributable to class origin. Note that it is also implied that the origin and destination linear effects are of the same sign (or one of them is zero), and that the size of each linear effect is constrained to lie between 0 and $\alpha + \beta$, inclusive.

Assuming these proportionality constraints are valid, multiplying the origin weight by the main diagonals will give the following set of origin effects:

$$\mu_{i[j=i]} w_{\text{o}} = \mu w_o + [(\alpha + \beta) w_o](i - i^*) + \widetilde{\alpha}_i, \tag{11}$$

where w_o is defined as in Equation 9. Because $w_o = \mu_\alpha/\mu$ as well as $\alpha/(\alpha+\beta)$, plugging in these values for the origin weight into Equation 9 will recover the values of μ_α and α, respectively. Likewise, multiplying the destination weight by the main diagonals will give the following set of destination effects:

$$\mu_{[i=j]j} w_d = \mu w_d + [(\alpha+\beta)w_d](j-j^*) + \widetilde{\beta}_j, \tag{12}$$

where w_d is defined as in Equation 10. Similar to the above, because $w_d = \mu_\beta/\mu$ as well as $\beta/(\alpha+\beta)$, substituting these values for the destination weight into Equation 10 will recover result in the values of μ_β and β, respectively.

Under the proportionality constraint, simple formulas for calculating the origin, destination, and mobility linear effects can be derived. Specifically, given the proportionality constraint, the origin linear effect generated by the DRM is $\widehat{\alpha} = (\alpha+\beta)w_o$, and the estimated destination linear effect is $\widehat{\beta} = (\alpha+\beta)w_d$.[8] As well, the estimated mobility linear effect is given by $\widehat{\gamma} = (\gamma-\alpha)+(\alpha+\beta)w_o$ or, equivalently, $\widehat{\gamma} = (\gamma+\beta)-(\alpha+\beta)w_d$.[9] However, again it should be emphasized that these calculations assume that the proportionality constraint holds or, equivalently, that the weights are correct. Because the origin, destination, and mobility linear effects are not identified, the proportionality assumption with respect to the linear effects cannot be tested against the data: the plausibility of this assumption can only be justified by appealing to theory or substantive knowledge about the underlying processes of social mobility.

With actual data, the origin and destination weights estimated by the DRM (i.e., $\widehat{w}_o$ and $\widehat{w}_d$, respectively) are a function of the underlying origin and destination nonlinear effects. Intuitively this makes sense, as these are identifiable components of the model while the linear effects are not. If the actual origin and destination nonlinear effects conform to the proportionality constraint, then the DRM's estimated weights will correctly recover these underlying nonlinear effects.[10] If there are no underlying origin, destination, and mobility nonlinear effects, then the DRM will generate

[8] Note that, accordingly, the DRM assumes that the origin and destination linear effects have the same sign (or one of the linear effects is zero). Likewise, the size of each linear effect is bounded between zero and $(\alpha+\beta)$. We build on this information later to generalize the DRM to a bounding approach grounded in partial rather than point identification.

[9] Note that all of these joint slopes are identified. We discuss this issue further in a later section.

[10] If the actual nonlinear effects do not conform to the proportionality constraint, then the DRM will still enforce the constraint on the estimated nonlinear effects, essentially assuming that these are equal to the true nonlinear effects. However, this will not in general be the case.

origin and destination weights of $\widehat{w}_o = 0.500$ and $\widehat{w}_d = 0.500$. In other words, absent additional information about the origin and destination linear effects, the DRM will assign equal value to the origin and destination categories with respect to the main diagonals of the mobility table. This is not an unreasonable assumption if there is no additional information available about the magnitude or direction of the linear effects, but researchers should be aware that it is an assumption nonetheless, and not an intrinsic feature of the data. We revisit this issue later when we discuss a generalization of bounding analyses for mobility effects models.

The above suggests a procedure for at least partially testing the plausibility of the proportionality constraint against the data. By fitting the L-ODM under a constraint, such as a zero destination linear effect, we can obtain a set of estimates of the underlying overall nonlinear effects of origin, destination, and mobility. We can then compare these nonlinear effects with those estimated by the DRM. If these are discrepant, then this suggests that the DRM is inappropriate, inasmuch it fails to recover the actual nonlinear effects. Alternatively, we can test whether or not the relative ratios of the origin and destination nonlinear effects are the same, as required by the DRM's proportionality constraint (see Equations 9 and 10). For example, we can examine whether or not the ratio of the first origin nonlinear effect to the sum of the first origin and destination nonlinear effects is the same as the ratio of the second origin nonlinear effect to the sum of the second origin and destination nonlinear effects. If these ratios are the same, then this lends indirect support for the plausibility of the proportionality constraint. However, as we show below, even when the proportionality constraint holds with respect to the nonlinear effects, there is no guarantee that the DRM will recover the underlying mobility linear effect.

2. Sobel's Fertility Findings Revisited

In this section we re-examine Sobel's (1981) findings on fertility using the DRM. We show that, even though the DRM assumes that there are virtually no mobility linear effects, the data are consistent with very large negative and positive linear effects. Sobel's original data on fertility is shown in Table 2. The outcome is the average number of children ever born by father's occupation (origin) and husband's 1962 occupation (destination) among wives aged 42 to 61 years in March 1962 who were currently living with their husband in the OCG sample. The total sample size is $R = 5,958$.

We first examine the plausibility of the proportionality constraint assumed by the DRM in the case of Sobel's fertility data. As noted in the previous section, this can be partially tested against the data because the nonlinear effects are identified. We first fit the L-ODM model by fixing the destination slope to zero, which allows us to estimate the nonlinear effects. These estimates, which are expressed as coefficients for orthogonal polynomials,[11] are then converted to deviations from the grand (or overall) mean. We next fit the DRM and converted the estimated origin and destination effects into deviations orthogonal to their respective overall levels and linear components.[12] Finally, for both models we calculated the estimated relative contributions of the origin and destination nonlinear effects. Specifically, for each model and for each of the origin nonlinear effects we calculated $\widetilde{\alpha}_i/(\widetilde{\alpha}_i + \widetilde{\beta}_{[j=i]})$, while for each of the destination nonlinear effects we calculated as $\widetilde{\beta}_j/(\widetilde{\alpha}_{[i=j]} + \widetilde{\beta}_j)$. For example, the ith origin nonlinear effect is divided by the sum of the ith and jth origin and destination nonlinear effects.

Table 3 highlights the main findings with respect to our evaluation of the validity of the proportionality constraint. First, the nonlinear effects for origin and destination differ between the DRM and L-ODM model, in some cases substantially. On balance, however, in all cases the direction of the nonlinear effects is correctly captured by the DRM. Second, as shown in the second row of Table 3, with the DRM all of the ratios are the same for the origin and destination nonlinear effects. This reflects the fact that the estimated origin and destination weights for the DRM are $w_o = 0.361$ and $w_d = 0.639$, respectively. Lastly, as shown in the bottom row of Table 3, the weights needed to recover the actual nonlinear effects, as represented by the estimates using the L-ODM model, vary considerably across origin and destination categories. This further suggests that the proportionality constraint of the DRM is likely inappropriate in this case.

An additional issue is that even when the proportionality constraint of the nonlinear effects is

[11] Specifically, the coefficients for the L-ODM model with the destination slope fixed to zero are as follows: $\mu = 2.318$, $\alpha = -0.300$, $\alpha^2 = 0.047$, $\alpha^3 = -0.020$, $\alpha^4 = 0.008$, $\beta^2 = 0.058$, $\beta^3 = 0.042$, $\beta^4 = 0.031$, $\gamma_L = -0.208$, $\gamma^2 = -0.001$, $\gamma^3 = 0.009$, $\gamma^4 = 0.001$, $\gamma^5 < 0.001$, $\gamma^6 = -0.003$, $\gamma^7 = 0.003$, $\gamma^8 = -0.001$.

[12] Specifically, regarding the origin deviations (or nonlinearities), we first calculated each origin overall effect by multiplying the origin weight by the estimates of the main diagonal. Let $\boldsymbol{\alpha}$ denote an $I \times 1$ column vector of the DRM's origin effects and $\mathbf{O}$ an $I \times I$ matrix of a column of 1's and a set of $(I-1)$ orthogonal polynomials. To obtain the distinct linear and nonlinear effects as well as the contribution to the overall intercept, we calculated $\boldsymbol{\alpha}^{\perp} = (\mathbf{O}^T\mathbf{O})^{-1}\mathbf{O}^T\boldsymbol{\alpha}$. We then dropped the parameters for the overall origin level and linear effect, giving us a column vector $\widetilde{\boldsymbol{\alpha}}^{\perp}$ of dimension $(I-2)\times 1$. Let $\widetilde{\mathbf{O}}$ denote the matrix $\mathbf{O}$ with the column of 1's and linear component dropped, such that it is of dimension $I \times (I-2)$. As a final step, we calculated $\widetilde{\mathbf{O}}\widetilde{\boldsymbol{\alpha}}^{\perp} = \widetilde{\boldsymbol{\alpha}}$, which resulted in an $I \times 1$ column vector of the DRM's nonlinear deviations.

valid, the DRM can generate highly biased estimates of the underlying mobility linear effect. This is the case even if the underlying mobility linear effect is extremely large. To show this, we conduct simulations based on the DRM and fertility data. Specifically, we fit the DRM and calculate the underlying estimated coefficients as linear and nonlinear effects, similar to the L-ODM. Then we use these to set up the data generating parameters (DGP) in the simulations. Specifically, we first fix the origin, destination, and mobility linear effects to the values assumed by the DRM, which are $\alpha = -0.115$, $\beta = -0.203$, and $\gamma = -0.104$. Then, we varied the values of mobility linear effect from -1.000 to $+1.000$, while calculating values of the origin and destination linear effects that are consistent with the original fertility data.[13] The remaining data generating parameters are given by $\mu = 2.3171$, $\alpha^2 = 0.0411$, $\alpha^3 = 0.0136$, $\alpha^4 = 0.0161$, $\beta^2 = 0.0730$, $\beta^3 = 0.0242$, $\beta^4 = 0.0285$, $\gamma^2 = -0.0012$, $\gamma^3 = -0.0074$, $\gamma^4 = 0.0004$, $\gamma^5 = 0.0027$, $\gamma^6 = -0.0022$, $\gamma^7 = 0.0017$, $\gamma^8 = -0.0008$. In these simulations, like the Sobel data, we set the number of origin and destination groups at $I = 5$ and $J = 5$, respectively. We also replicated each combination of origin and destination (and accordingly mobility) groups so that number of individuals in each cell is identical to that of the original data. This resulted accordingly in a total sample size of $R = 5,958$ for each simulation. For simplicity, and without loss of generality, we assume no random error.

The results from the simulations based on Sobel's data are shown in Table 4. The table reveals that the DRM fixes the mobility linear effect to a particular value, $\hat{\gamma} = -0.010$, even in cases in which the mobility linear effect is quite large (either negative or positive). The shaded column in Table 4 indicates that particular simulation in which the DRM recovers the true mobility linear effect. The results from this table underscores that the DRM can generate results that are, to a great degree, inconsistent with the true mobility effect. Note that in all simulated data sets the DRM correctly recovers the underlying nonlinear effects. This is by construction, as we have purposely created the data so that the DRM's proportionality constraint on the nonlinear effects is valid.

[13] Using the DRM's estimates, we first calculated the sum of the origin and destination linear effects as well as the sum of the mobility and destination linear effects, both of which are identifiable. This gave us $\alpha + \beta = -0.317$ and $\gamma + \beta = -0.213$. We then calculated the corresponding origin and destination linear effects using these sums. Specifically, using the DRM estimates, the corresponding origin and destination linear effects for a given mobility linear effect γ are $\alpha = (\alpha + \beta) - (\gamma + \beta) + \gamma = -0.104 + \gamma$ and $\beta = (\gamma + \beta) - \gamma = -0.213 - \gamma$.

3. The Nature of the Bias from the DRM

In the previous example, why does the DRM fix the mobility slope to nearly zero even when the mobility linear effect is quite large? When the underlying nonlinear effects obey the proportionality constraint, the nature of the bias from the DRM can be expressed as a mathematical function of the estimated weights. Specifically, assuming the true nonlinear effects conform to the proportionality constraint (see Equations 9 and 10), the bias in the DRM's estimated mobility linear effect can be written as[14]

$$\hat{\gamma} = \text{ True Linear Effect} + \text{Bias} = \gamma + [\beta(\widehat{w}_o) - \alpha(\widehat{w}_d)]. \tag{13}$$

This equation clarifies that the DRM estimate will only be unbiased if $\beta(\widehat{w}_o) = \alpha(\widehat{w}_d)$. Because the origin and destination linear effects are not identifiable from the data, this is a constraint that can only be justified by appealing to sociological theory or substantive knowledge external to the data at hand. In general, this equation clarifies that the bias in the estimated mobility effect arises from two main sources: first, the size and sign of the underlying linear effects α and β; second, the estimated weights $\widehat{w}_o$ and $\widehat{w}_d$, which are in turn affected by the size and sign of the nonlinear effects. Note that if the origin and destination nonlinearities are zero, then the DRM will fix $w_o = w_d = 0.500$.

To illustrate the sensitivity of the DRM, we first show the bias in the mobility linear effect across different values of the true origin and destination linear effects. In these simulations we set the nonlinear effects to zero, so the DRM's estimated origin and destination weights $\widehat{w}_o = 0.500$ and $\widehat{w}_d = 0.500$, respectively. These results are shown in Table 5.[15] The bias formula for the mobility linear effect is displayed on the right half of the table, while the left half shows the true linear effects as well as those produced by the DRM. The top half of the table displays the simulations for differing values of the origin linear effect while keeping the values of the other parameters in the data generating model the same. The shaded row indicates that particular simulation in which the DRM recovers the true mobility linear effect. As can be seen from this table, the mobility linear effect is

[14] Likewise, again assuming that the proportionality constraint holds with respect to the nonlinear effects, the bias of the estimated origin linear effect is given by $\alpha^* = \alpha + [(\alpha + \beta)\widehat{w}_d - \beta]$, and the bias of the estimated destination linear effect is given by $\beta^* = \beta + [(\alpha + \beta)\widehat{w}_o - \alpha]$.

[15] By varying the size and sign of the true origin and destination linear effects, we are altering the true weights needed to recover the underlying origin and destination linear effects. The proportionality constraint is thus violated for the linear effects. However, this does not affect the estimated weights of the DRM, which, as noted in the main text, simply gives an estimated weight of 0.500 for both origin and destination when there are no underlying nonlinear effects.

easily biased due to differing values of the origin linear effect. The bottom half of the table shows a similar set of simulations, this time varying the values of the destination linear effect. Again, the mobility linear effect is biased for all rows except that one which is shaded. In summary, the simulations in Table 5 show that, under plausible values of the data generating models, the DRM can easily produce biased estimates of mobility effects.

The bias of the DRM's estimated mobility linear effect depends not only on the true origin and destination and linear effects, but also the underlying nonlinear effects. That is, because the bias formula depends on the DRM's estimated weights, the estimated mobility linear effect from a DRM is in turn a function of the underlying nonlinear origin and destination effects.[16] In Online Appendix E, Table E.1 we present various results from simulations with various values of the nonlinear effects for origin and destination. Together, these simulations reveal that the estimated mobility effects are quite sensitive to the true values of the nonlinear origin and destination effects.

In short, DRM relies on a specific proportionality constraint to identify unique origin, destination, and mobility effects in a mobility table. Unfortunately, because the linear effects are not identified, the validity of this assumption can only be tested indirectly using the nonlinear effects. Even more problematic, regardless of the size of the underlying linear mobility effect, the DRM will fix the linear effect to a very specific value. For example, as illustrated using Sobel's (1981) fertility data, even if the mobility linear effect is extremely positive or negative, the DRM still fixes the mobility linear effect close to zero. Finally, we have provided a simple bias formula for the mobility linear, showing that the estimates from the DRM are sensitive to both the size and sign of the underlying origin and destination linear effects, as well as to the nonlinear effects. Taken together, these results suggest that the DRM should not be used to identify the effects of social mobility unless its very specific assumptions are strongly supported by sociological theory or background knowledge.

While the conventional DRM assumes fixed origin and destination weights, alternative specifications have been proposed that allow these weights to vary across the mobility table. Sobel (1985) extended the model to permit weights to differ by either origin or destination class, while Weakliem (1992) further allowed weights to vary simultaneously by both origin and destination. These

[16]Note again that we are assuming that the underlying nonlinear effects conform to the proportionality constraints shown in Equations 9 and 10. If the actual nonlinear effects do not conform to these constraints, then the bias formulas are invalid as the nonlinear effects cannot be described by a single weight.

extensions can accommodate more complex patterns of nonlinear effects in the data. Nevertheless, a fundamental limitation persists: due to the inherent identification problem, these extended DRM variants still impose specific constraints to identify unique linear effects of origin, destination, and mobility. In essence, all DRM formulations, regardless of their flexibility in handling nonlinearities, ultimately rely on point identification, requiring researchers to make highly specific and, as a consequence, strong assumptions about the magnitude and direction of the linear effects. Crucially, as we demonstrate in subsequent sections, the core assumptions of mobility effects models cannot be directly tested against the empirical data, making any results contingent in a nontrivial way on untestable modeling choices.

II. Generalizing Models of Mobility Effects

In this section, we show how previous generations of mobility effects models, such as the SAM and the DRM, can be understood as special cases of a bounding analysis. We first present a visualization of the identification problem at the core of mobility effects models, showing how the DRM and SAM achieve point identification by imposing extremely strong assumptions on the magnitude and direction of the linear effects. We then show how one can construct bounds on the linear effects using a variety of other constraints that generally involve much weaker assumptions. It is important to emphasize, however, that the results of any analysis of mobility effects are only as valid as the social theory or substantive knowledge used to justify the bounds.

1. Point Identification and the Canonical Solution Line

To understand how the DRM, SAM, and related models can be viewed as special cases of a bounding approach, it is crucial to recognize the geometric interpretation of the non-identifiability of the linear effects. Because mobility effects models are not fully identified, we cannot obtain point estimates for each of the origin, destination, and mobility linear effects (or effects that are partially a function of the linear effects). A convenient way to express the identification problem is to note that for any particular mobility effects model we can specify the slope as:[17]

[17] Following our previous discussion, we will assume without loss of generality that we have applied sum-to-zero constraints.

$$\alpha^* = \alpha + \nu, \;\; \beta^* = \beta - \nu, \text{ and } \gamma^* = \gamma + \nu, \tag{14}$$

where the asterisk ($*$) indicates an arbitrary set of estimated slopes from a mobility effects model under some particular constraint and ν is a scalar fixed to some value. As Equation 14 indicates, the estimated parameters are simple transformations of the true unobserved slopes α, β, and γ shifted by a single arbitrary scalar, ν. Importantly, selecting a value of the scalar ν is equivalent to specifying a particular constraint to identify a mobility effects model.

By varying values of ν (which can take on any value), we trace out what has been called the *canonical solution line* in a parameter space defined by the range of possible origin, destination, and mobility slopes (Fosse and Winship 2018, 2019b). It is a "solution" line because any set of possible estimates of the linear effects will lie on this line, and it is "canonical" because it is the simplest possible geometric representation of the identification problem.[18] Note that, if a mobility effects model were identified, then there would be not a line but a single point in the parameter space.

Figure 2 shows the canonical solution line based on the Sobel (1981) fertility data.[19] With Sobel's fertility data, $\widehat{\Gamma}_1 = -0.317$ and $\widehat{\Gamma}_2 = -0.213$. Several crucial points are worth emphasizing regarding this figure. First, in the absence of data from a mobility table and a corresponding mobility effects model, the origin, destination, and mobility slopes may take on any combination of values in a three-dimensional parameter space. With a mobility effects model, vast areas of the parameter space can be ruled as inconsistent with the data. However, as we discuss later, this does not mean that quite strong assumptions are necessary for identifying unique mobility effects. Second, depending on the data, the location of the canonical solution line relative to the origin will differ, leading to different trade-offs from setting various constraints. Specifically, the location of the canonical solution line is a function of the parameters $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$. With So-

[18] More specifically, given a linear dependency among origin, destination, and mobility, other design matrices, such as those based on treatment, sum-to-zero deviation, or Helmert contrasts, will have a solution line lying in a q-dimensional parameter space, where q is greater than three. However, as Fosse and Winship (2018) show, one can construct a transformation matrix that will convert the solution line of any particular design matrix into a canonical form that lies in just three dimensions defined by the range of possible parameter values for origin, destination, and mobility.

[19] To keep our findings consistent with the previous discussion on the DRM, which assumes that the true nonlinear effects conform to the proportionality constraint, we will use the simulated Sobel (1981) data. Our main conclusions hold for actual data, except that, because the DRM's proportionality constraint is violated, the linear effects generated by the DRM will not lie exactly on the canonical solution line (see Figure 3). For the remainder of the article, for simplicity we will refer to Sobel's data, but this specific usage should be kept in mind.

bel's fertility data, these values are $\widehat{\Gamma_1} = -0.317$ and $\widehat{\Gamma}_2 = -0.213$. Note that these parameters are identified (as well as their difference) because the scalar ν in Equation 14 cancels out. For example, $\alpha^* + \beta^* = \alpha + \nu + \beta - \nu = \Gamma_1$.[20] Finally, because the canonical solution line is a function of just two parameters (Γ_1 and Γ_2), we can reduce our three-dimensional representation to just two dimensions. One way of doing this is by having the horizontal axis represent the destination slope, the left vertical axis the origin slope, and the right vertical axis the mobility slope (see Fosse and Winship 2018).

Figure 3 shows a two-dimensional representation of the canonical solution line shown in Figure 2, again using Sobel's (1981) fertility data. As in Figure 2, this line is based on values of $\widehat{\Gamma}_1 = -0.317$ and $\widehat{\Gamma}_2 = -0.213$. It is important to note that the solution line shown in Figure 3 is identical to the one shown in Figure 2. Each location in the coordinate space can be referenced in terms of the origin, destination, and mobility linear effects, so that, for example, the point $(-0.167, -0.150, -.063)$ refers to $\alpha = -0.167$, $\beta = -0.150$, and $\gamma = -0.063$.[21] The solution line will always run from top left to bottom right, and the slope of the solution line relating destination to origin and mobility will always be -1. Not all features of the solution line, however, are invariant. As noted above, the values of Γ_1 and Γ_2 determine two crucial features of the canonical solution line. First, the difference between Γ_2 and Γ_1, estimated to be 0.104 in Figure 3, determines the offset between the origin and mobility scales. Thus, in Figure 3, when $\gamma = 0$, $\alpha = -0.104$. Second, at the point where $\beta = 0$, Γ_1 and Γ_2, respectively, determine the location of the canonical solution line with respect to α and γ. Various traditional mobility effects models can be located on the solution line in Figure 2, and can be understood as making very specific assumptions about the underlying origin, destination, and mobility linear effects.

Three estimates are displayed in Figure 3. First, as shown by the blue triangle in Figure 3, there is the set of estimates corresponding to Duncan's SAM. As noted previously, the SAM's estimates are equivalent to assuming that the mobility linear effect is zero (i.e., $\gamma = 0$).[22] Second, there are

[20] More generally, for any arbitrary pair of real numbers (p, q), the linear function $p\alpha + (p+q)\beta + q\gamma$ can be estimated from the data. To see this, note that $p\alpha^* + (p+q)\beta^* + q\gamma^* = p(\alpha + v) + (p+q)(\beta - v) + q(\gamma + v) = p\alpha + pv + p\beta - pv + q\beta - qv + q\gamma + qv = p\alpha + (p+q)\beta + q\gamma$.

[21] This assumes that this particular point reflects the underlying origin, destination, and mobility linear effects. One could refer to the location more generally as α^*, β^*, and γ^*, which reflects a set of constrained values that is not necessarily equal to the true, underlying linear effects.

[22] Note, however, that in practice the SAM will not exactly equal a zero linear effect because of some bias due to the exclusion of the mobility nonlinear effects from the model (for details, see Equation C.11 in Online Appendix C).

the estimates for the DRM, as shown by the green circle in Figure 3. As noted previously, assuming that the nonlinear effects satisfy the proportionality constraint, the mobility linear effect is fixed to $(\Gamma_2 - \Gamma_1) + \Gamma_1 w_o$ or, equivalently, $\gamma = \Gamma_2 - \Gamma_1 w_d$, where again w_o and w_d are the origin and destination weights, with $w_d = 1 - w_o$. In the case of Sobel's fertility data, assuming the proportionality constraint is valid for the nonlinear effects, then the DRM will fix the mobility linear effect to -0.010 (see Table 4).[23]

The DRM and the SAM have traditionally been the most popular ways of point identifying mobility effects, but these are not the only options. Making any assumption about the sign and size of one of the three linear effects is sufficient to point identify the remaining two linear effects. Thus, for example, one might assume that the origin linear effect is zero, or that the destination linear effect is some specific negative value. One particularly attractive approach is to use what we call the *same-slopes assumption*. Presumably, origin and destination effects reflect similar, if not identical, underlying causal processes. Accordingly, in some contexts it might be reasonable, at least as a first pass, to assume that the origin and destination linear effects are the same, such that they have the same direction and magnitude. In other words, we might assume that origin and destination linear effects are the same such that $\alpha = \beta$. This will then provide an estimate of the underlying linear mobility effect, or γ. This constraint is easily derived from the data. Because $\Gamma_1 = \alpha + \beta$ is identified, under the same slopes assumption we know that $\alpha = \Gamma_1/2$ and $\beta = \Gamma_1/2$. We also know that $\Gamma_2 = \gamma + \beta$, so plugging in for $\beta = \Gamma_1/2$ and rearranging terms we have $\widehat{\gamma} = \Gamma_2 - \Gamma_1/2$. This is the mobility linear effect under the same-slope assumption. Figure 3 shows the point estimates under the same-slope assumption as a red square. Under the same-slopes assumption, the parameter Γ_1 is split equally between origin and destination. With Sobel's fertility data, $\widehat{\Gamma_1} = -0.317$ and $\widehat{\Gamma}_2 = -0.213$. Thus, given the same-slopes assumption, we have $\widehat{\alpha} = -0.159$ and $\widehat{\beta} = -0.159$ and, accordingly, $\widehat{\gamma} = -0.055$.

Each of these three estimates of the mobility linear effects correspond to different patterns of overall mobility effects, which incorporate the nonlinearities. Figure 4 shows the estimated mo-

Because of this bias, in some cases the SAM's estimates may not lie exactly on the canonical solution line. An alternative is to fit the Diff-SI model (see Equation C.2 in Online Appendix C). Under the assumption that the mobility linear effect is zero, the slope parameters for origin and destination will be similar to those of the SAM but without the bias that results from excluding the mobility nonlinear effects.

[23] However, if the proportionality constraint is not satisfied with respect to the nonlinear effects, then the DRM's estimates may not lie exactly on the canonical solution line.

bility effects using the three identification strategies discussed above. The main conclusion is that both the DRM and the SAM, as indicated by the blue and green lines in Figure 4, are broadly consistent with no meaningful pattern of mobility effects, apart from the up-and-down pattern of the nonlinear effects. By contrast, the assumption of equal origin and destination slopes yields an overall negative effect of social mobility on fertility, as indicated by the red line in Figure 4. That is, downward (or upward) social mobility causes households to have fewer (or more) children.

The results in Figure 4 are compelling, but great caution is warranted. Only the nonlinear effects in Figure 4 are identified, while the linear effects are a function of the particular model employed, which encodes very specific assumptions about the size and sign of the linear effects. Because there is an infinite number of possible linear effects that are consistent with the data, there is also an infinite number of possible results that could be derived by applying any of a variety of mobility effects model (including those that have not yet been devised). For example, constructing a model that assumes that the origin linear effect is very negative will result in a very negative mobility linear effect, and thus a steep downward pattern in Figure 4. Conversely, using a model that assumes the origin linear effect is positive will result in a very positive mobility linear effect and, accordingly, a steep upward pattern in Figure 4. The extreme sensitivity of mobility effects models to the results obtained has not been generally appreciated in the literature. It should also be emphasized that the conventional mobility effects models rely on point identification, which requires invoking very specific and extremely strong assumptions that are not directly testable against the data. In the next section, we outline a number of identification strategies that involve weaker assumptions, albeit at the potential expense of the precision of the estimates.

2. Partial Identification with Bounding Analyses

Conventional models of mobility effects, as outlined above, are based on point identification, where the true parameter is uniquely estimated from the data given the specification of a particular model. Unfortunately, in the case of mobility effects, point identification depends crucially on the model (or, equivalently, on assumptions about the linear effects). As we have shown, different models (or different assumptions) lead to different estimates of mobility effects. A more principled approach is to abandon point identification in favor of partial identification. Rather than trying to extract a single estimate of a linear mobility effect, one uses (potentially weaker) assumptions to generate a range of estimates, thus partially identifying (or bounding) the mobility effect.

In the case of point identification, fixing the values of one of the slopes determines the values of the other two. For example, assuming that the mobility linear effect is zero, as is the case with the SAM, will fix the values of the origin and destination linear effects. Similarly, setting upper and lower bounds on the magnitude and direction of any one of the linear effects will automatically set bounds on the size and sign of the remaining two linear effects. Likewise, setting the sign of any two linear effects will establish bounds on the size and sign of the remaining linear effect.[24] These two strategies for specifying bounds are based only on assumptions about the magnitude and direction of the underlying linear effects. However, a third strategy, which may entail even weaker assumptions, is to use the nonlinear effects to make assumptions about the shape of one or more of the effects over a particular range of the data. For example, one might assume that the pattern of destination effects is monotonically increasing, decreasing, or neither increasing nor decreasing for some (possibly restricted) set of destination categories. These assumptions will, in turn, place constraints on the underlying linear effects.[25] Finally, as Manski (1990, 2003) already pointed out in his seminal work on partial identification, there is a direct trade-off between the strength of one's assumptions and the width of the bounds on the parameters. While, in some applications, analysts may bemoan the width of the bounds to be too large to be analytically useful, wide bounds merely demonstrate that much stronger theoretical assumptions are required if the analyst seeks to usefully interpret a given estimate.

To illustrate how one might proceed with a bounding analysis of mobility effects, we outline three main partial identification strategies. First, there is what we call the *same sign assumption*.

[24] Formulas for setting bounds on the size and sign of one slope, as well as setting the sign of two linear slopes, are shown in Tables A.1 and A.2, respectively, in Online Appendix A.

[25] For example, suppose that we have strong theoretical reasons to believe that the overall destination effect is monotonically increasing. Our task, then, is to specify a value for the destination linear effect that ensures that the total effect (which includes the nonlinear effects in addition to the linear effect) is monotonically increasing. This implies that between any two adjacent destination categories, the pattern of effects is either flat or increasing (but not decreasing). To determine the minimum slope needed for a monotonically increasing set of effects, we simply find the pair of adjacent destination categories for which the downward pattern is most negative. For example, suppose the forward differences for the destination nonlinearities are $\Delta\widetilde{\beta}_{I-1} = \{1, 3, -1, -1, -2, 1.5, -1.5\}$. The minimum of these differences is -2. To counteract this downward deviation, the parameter value for the destination slope must be greater than or equal to $+2$ so as to ensure that the overall pattern is monotonically increasing. Similarly, one can derive slopes based on the assumption that the overall pattern of effects is monotonically increasing or neither monotonically increasing nor decreasing over some range of the data (for examples, see Fosse and Winship 2019b; Gowen et al. 2023).

Rather than assuming that the origin and destination linear effects have the same magnitude and direction, as with the same-slopes assumption, in some settings it might be appealing to assume that origin and destination have the same sign but not necessarily the same size. The justification for this assumption is similar to that for the same-slopes assumption, namely, that the causal processes proxied by class origin and destination are similar, if not identical.

The constraints implied by the same sign assumption are easily calculated from the data. Because $\Gamma_1 = \alpha + \beta$ is identified, under this assumption we know that α and β each range from 0 (lower bound) to Γ_1 (upper bound). We also know that $\Gamma_2 = \gamma + \beta$, so plugging in for $\beta = 0$ (lower bound) and $\beta = \Gamma_1$ we have $\widehat{\gamma} = (\Gamma_2 - \Gamma_1, \Gamma_2)$. In the case of Sobel's fertility data, the same-sign assumption results in bounds of $\widehat{\alpha} = (-0.312, 0)$, $\widehat{\beta} = (-0.312, 0)$, and $\widehat{\gamma} = (-0.213, 0.104)$. These bounds and the resultant set of estimates on the canonical solution line are shown in panel (a) of Figure 5. The red rectangle indicates the range of estimates compatible with the assumption about the origin slope, while the blue rectangle denotes the set of estimates compatible with the assumption about the destination slope. Their intersection on the canonical solution line, which shows the set of estimates compatible with the data, is shown as a bold line. Unfortunately, in this particular case these bounds are not particularly informative, as the slope for the mobility linear effect may be both positive, zero, or negative. This is evident in Figure 6(a), which shows the resultant overall mobility effects under the same-sign assumption,

A second strategy that one might employ with mobility data is what we call the *monotonic origin-destination effects assumption*. The idea here is that, again to the extent that origin and destination can be understood as having similar underlying causal processes, we might assume not only that both of their linear effects have the same sign, but that that one or both of the effects are monotonically increasing (if Γ_1 is positive) or monotonically decreasing (if Γ_1 is negative).

Some care is warranted, however, in calculating underlying linear effects based on monotonicity constraints. In many if not all cases, one might believe that the origin and destination nonlinear effects capture, at least in part, some amount of random error. Applying monotonicity constraints in such circumstances may thus be misleading. An alternative, which we use here, is to calculate the monotonicity constraints based only on some restricted set of higher-order terms, such as only quadratic component or only the quadratic and cubic components. In other words, the effect is assumed to be monotonic only with respect to these lower-order terms, while nonlinear effects may themselves be monotonically increasing, randomly fluctuating, or monotonically decreasing.[26] For Sobel's fertility data, we calculate the monotonicity constraints using only the quadratic component, but substantively similar findings were obtained using higher-order terms.[27]

Based on the quadratic linear component, the maximum possible slope for a monotonically decreasing origin effect is -0.123, so we specify bounds on the origin linear effect as $\widehat{\alpha} = (-\infty, -0.123)$. For the destination linear effect, we keep the assumption that the slope is negative, such that $\widehat{\beta} = (-\infty, 0)$.[28] Combined, these bounds result in a mobility linear effect of $\widehat{\gamma} = (-0.206, -0.026)$. The bounds for the linear effects, as well as that part of the solution line consistent with these bounds, are shown in panel (b) of Figure 5. The overall mobility effects, which incorporate the (identified) mobility nonlinear terms, are shown in Figure 6(b)). Given these assumptions, downward (or upward) mobility causes households to have many fewer (or more) children. Specifically, those who the most upwardly mobile ($+4$) have approximately 1.11 fewer children than those who are the most downwardly mobile (-4).

So far we have employed two approaches to partially identifying mobility effects, one based on the same-sign assumption and another based on monotonically decreasing origin and/or destination effects. A third approach is to invoke the assumption of monotonically decreasing effects, but to do so for some restricted range of class categories. For example, suppose we assume that fertility rates across class origin categories are monotonically decreasing, but only those classes lower in the hierarchy (1, 2, and 3). This assumption aligns with the claim that fertility may decline consistently across individuals from lower class origins, as even modest differences in resources, education, health care, and family planning may significantly shape reproductive strategies. However, for in-

[26] An alternative approach is to adopt a fully Bayesian approach and "smooth" the data by using very strong priors on the higher-order polynomials, effectively shrinking them to zero, or nearly so. For an example, see Fosse (2021).

[27] Specifically, the bounds on the canonical solution line using all higher-order terms are narrower than those based only on the quadratic component and, as a consequence, the bounds on the overall mobility effects are also narrower. This is likely unrealistic, as we expect some of these higher-order nonlinear effects to be noise.

[28] If we assume that the destination effect is monotonically decreasing with respect to its quadratic component, then its maximum slope is -0.219. As a result, assuming that both the origin and destination linear effects are monotonically decreasing with respect to the quadratic component is incompatible with the data, as there is no part of the canonical solution line that is consistent with the assumptions that $\widehat{\alpha} = (-\infty, -0.123)$ and $\widehat{\beta} = (-\infty, -0.219)$. In other words, any theory that makes the claim that the effects of both origin and destination are monotonically decreasing with the quadratic component can be shown to be false. This is not an artifact of just using the quadratic component in calculating the monotonicity constraints. Additional analyses reveal that with the full set of nonlinear effects, the data are inconsistent with the assumption that class origin and destination are both monotonically decreasing.

dividuals from middle class origins or higher, the primary determinants of lower fertility might be largely satisfied, resulting in stabilized fertility levels across these class groups, or at least fertility levels that are not monotonically decreasing.

The assumption that fertility is monotonically decreasing, but only across class origin categories lower in the hierarchy, results in a maximum possible origin slope of -0.0922. Accordingly, we specify the bounds on the origin linear effect as $\widehat{\alpha} = (-\infty, -0.0922)$. For the destination linear effect, we again keep the assumption that the slope is negative, such that $\widehat{\beta} = (-\infty, 0)$. Together, these bounds yield a mobility linear effect of $\widehat{\gamma} = (-0.206, -0.004)$. The bounds for the linear effects, as well as that part of the solution line consistent with these bounds, are shown in panel (c) of Figure 5, while the overall mobility effects are shown in Figure 6(c). Under these assumptions, we arrive at qualitatively the same conclusions as in Figures 5(b) and 6(b). That is, we can conclude that those who the most upwardly mobile (+4) have approximately 0.99 fewer children than those who are the most downwardly mobile (-4).

Alternatively, suppose we assume that fertility rates across class origin categories are monotonically decreasing, but only those classes higher in the hierarchy (3, 4, and 5). This assumption is consistent with the claim that fertility differences are most pronounced among individuals from higher class origins, as distinctions in economic resources, aspirations, and family formation norms become increasingly influential at higher class origin positions. By contrast, among lower class origins, fertility rates may already reflect uniformly constrained resources and limited opportunities, resulting in less variation in fertility across these categories. The assumption that fertility is monotonically decreasing, but only across class origin categories higher in the hierarchy, yields a maximum possible origin slope of -0.244 and thus bounds on the origin linear effect of $\widehat{\alpha} = (-\infty, -0.244)$. As in the prior examples, for the destination linear effect we assume that the slope is negative, such that $\widehat{\beta} = (-\infty, 0)$. These bounds together result in a mobility linear effect of $\widehat{\gamma} = (-0.206, -0.146)$. Figure 5 (d) shows the bounds on the linear effects, including that section of the solution line that is consistent with these bounds. The overall mobility effects are shown in Figure 6(d), which have quite narrow lower and upper bounds. The main conclusion, in line with the results in panels (b) and (c), is that downward mobility has a strong negative effect on fertility. Specifically, we can conclude that those who the most upwardly mobile (+4) have about 1.59 fewer children than those who are the most downwardly mobile (-4).

The above approaches to identifying mobility effects employ much weaker assumptions than conventional models, such as the DRM and SAM, which rely on point identification. This is not to imply that these assumptions are not very strong, however.[29] In many applications, theories or additional information about the underlying causal processes may be quite ambiguous, precluding any practical guidance on how to set bounds on the effects. Moreover, even when one believes that theory or background information provides useful guidance on how to set the bounds, it is absolutely critical to understand that the results are by definition extremely sensitive to the assumptions invoked. This means that any analysis that attempts to identify unique effects of mobility is inherently provisional, and thus should always be presented transparently, clearly specifying the assumptions invoked and explicitly discussing the tentative nature of the conclusions. As we have shown, this is facilitated by parameterizing the model so that the unidentified part is separate from the unidentified part, as we have done with the L-ODM model.

Finally, it is worth underscoring that even if one believes that the true effects have been identified, there are additional and serious complications with treating such effects as "causal" in any meaningful sense. As Zang, Sobel, and Luo (2023) have recently emphasized, it is important to recognize that the DRM and related models were developed in a historical context that largely predated modern frameworks for causal inference. While these techniques represented innovative solutions to the methodological challenges of their time, they now face significant limitations when evaluated against contemporary standards for causal identification. In Online Appendix F, we outline in detail these issues regarding parallel-world counterfactuals, the consistency assumption, the assumption of no confounding given composite exposures, and complications that arise due to the fact that mobility is likely an exposure-induced confounder. These difficulties mean that even if the underlying bundles of mechanisms for origin, destination, and mobility could somehow be observed, there remain substantial barriers to interpreting estimated parameters as causal effects. These problems

[29]The substantive utility of the bounding approach, even with potentially wide bounds, is considerable. It improves methodological transparency by requiring explicit statement and justification of identifying assumptions, unlike methods where strong assumptions might be implicit or untestable. It also accurately reflects the range of conclusions supportable by the data under relatively weaker, theory-driven constraints, thus avoiding the potentially false precision of point estimates. Furthermore, the canonical solution line itself delineates all linear effect combinations consistent with the data, thus falsifying any combination not lying upon it. Finally, the specific identified set derived from bounding assumptions allows for the potential falsification of theoretical claims or previously reported point estimates if they fall outside the established bounds.

apply not only to the DRM and SAM but also to methods using bounds to partially identify effects.[30]

III. Paradigm Shift: Towards a Positional Sociology of Social Mobility

The prior sections have critically assessed the DRM and the general challenges for the identification of independent origin, destination, and mobility effects. Taken together, we believe that the problems identified call for a fundamentally new direction in the analysis of the consequences of social mobility. We now begin to outline what we consider a particularly promising way forward. To be clear, this entails a re-framing of the analytic question. But, in words typically ascribed to Charles Kettering, "a problem well stated is a problem half solved."

The alternative framework for analyzing social mobility that we propose in this section is based on the concept of *positional sociology* (Fosse and Pfeffer 2023). This approach differentiates between two distinct forms of inequality: (1) *structural inequalities*, associated with stable positions within the social hierarchy, and (2) *dynamic inequalities*, arising from movements between social positions. Building on this distinction, we introduce the Structural and Dynamic Inequality (SDI) model, which extends the L-ODM model by explicitly re-indexing parameters according to social origin and mobility status. The result is a fully identified model with clearly interpretable parameters that effectively capture both types of inequality. As we go on to show, this method enables a decomposition of population-level variation within mobility tables, leading to highly informative sets of parametric expressions and visualizations.

Our approach broadly aligns with recent contributions by Zang, Sobel, and Luo (2023) and Breen and Ermisch (2024). Like us, these authors emphasize challenges in isolating unique mobility effects and suggest alternative conceptualizations of social mobility. Specifically, they conceptualize class destinations as "treatments" whose effects vary across origin categories. However, we adopt a model that, while fully identified, nonetheless leverages the distinct, identifiable components of origin, destination, and mobility. Furthermore, and more importantly, our framework diverges from their approach by explicitly emphasizing mobility as a positional dimension rather than as a variation in destination treatment effects. Consequently, our primary objective is to use the SDI model to systematically dissect mobility tables into their constituent structural and dynamic inequalities,

[30]For a critique of an alternative approach to estimate mobility effects also see (Zhou/Song, in this issue).

thereby highlighting the inequalities experienced by distinct social groups.[31]

1. Mobility as a Metric of Positionality

In contrast to conventional models of mobility effects, what we term "positional sociology" takes a distinctly different approach to the deterministic relationship between origin, destination, and mobility (Fosse and Pfeffer 2023). Whereas conventional models regard this relationship as an obstacle to be overcome, positional sociology embraces it as a natural observational fact, recognizing that origin, destination, and mobility are interrelated dimensions that locate positions in an observed social space, rather than as proxies for distinct bundles of causal mechanisms that must be separated. Positional variables are not themselves causal, but are axes along which variation is observed. Treating them as causal in of themselves is a fundamental "category mistake" (e.g., see Ryle 1959).

Suppose, for example, that an individual from a given class origin group moves up the social hierarchy and reaches a new class destination. Reaching this new position in the social structure is observed conjointly with the degree of social mobility. From this perspective, social mobility as a dimension of positionality can never be completely separated from class destination, nor should it be, because social mobility is experienced in the context of class destination rather than independently of it.[32] However, treating class, origin, and mobility as dimensions of social position rather than as surrogates for underlying causal factors does not mean that positional sociology does not involve the study of causal processes; far from it. Rather, the causal factors are distinct and separate from the dimensions of social position, which describe the observed locations of individuals and other groups in an observed multidimensional social space.[33]

[31] Although our primary application of the SDI model focuses on summarizing structural and dynamic inequalities, the model's parameters can also be interpreted causally under conventional assumptions. Rather than identifying isolated effects of each dimension entirely, the SDI model identifies joint effects. For instance, instead of isolating the unique effect of origin alone, our model identifies the combined effect of origin and destination; similarly, rather than isolating mobility effects alone, it identifies the joint effect of mobility status and destination. It is important to emphasize, however, that identifying these joint causal effects necessitates standard but quite strong assumptions, notably those related to consistency and the absence of unobserved confounding.

[32] For a formal treatment of the identification problem discussed here, see Online Appendix B.

[33] To reiterate, in the following sections we focus on estimating joint sets of parameters. These can be treated as descriptive quantities, or, if stronger assumptions about consistency and unconfoundedness are invoked, as joint causal effects, which, as a careful inspection of Figure F.2 will show, are identified (assuming the validity of an additive model of the underlying causal factors). An alternative perspective, more in line with positional sociology, is that these quantities represent descriptive quantities, but can be informally treated as "causal" using a decomposition analysis (see, e.g., Fosse and Winship 2019a; Jackson and VanderWeele 2018).

Given this epistemology, a key question is: how exactly should we describe population-level variability using information from a mobility table? There are six logically possible ways, as shown in Table 6. Let $f(.)$ denote some generic function and, as before, let O, D, and M denote the observed origin, destination, and mobility dimensions on a mobility table. We can then describe variability with models of the form $f(O)$, $f(D)$, $f(M)$, or with models of the form $f(O,D)$, $f(O,M)$, $f(D,M)$. For example, researchers often describe differences by class destination (e.g., Goldthorpe 1999), which corresponds to a function of the form $f(D)$. Duncan's SAM can be re-interpreted as a model that does not attempt to extract unique origin, destination, and mobility effects, but rather as an $f(O,D)$ model that attempts to describe the surface of a mobility table (see also Online Appendix B).

Among these various generic models for summarizing population-level variability in a mobility table, we have a strong preference for models of the form $f(D)$ and $f(O,M)$, which are shaded gray in Table 6. The reason is that models based on the other functions, although quite common, conflate structural with dynamic inequalities and thus fail to distinguish social structure, or the "map of locations," from social mobility, or "movements from one location to another" (Fosse 2023).[34] We elaborate this view below. We also refer the reader to Online Appendix C, which contains detailed mathematical formulas outlining the limitations of the other models shown in Table 6.

Consider first a model of the form $Y = f(O,M) + \epsilon$, where ϵ is a normally distributed error term with a mean of zero. For some summary outcome Y observed in a given class destination $D = O + M$, what we will refer to as *social mobility analysis* is focused on examining models of the form $Y = f(O,M) + \epsilon$. Conditional on M, variability in Y across levels of class origin O (and thus also D) reveals a social structure differential, or structural inequality. This represents the "map of locations" with respect to the outcome being examined. Conversely, conditional on O, variability in Y across levels of class mobility M (and thus also D) reveals a social mobility gradient, or a dynamic inequality. This represents, for a given class origin, how the outcome is related to "movements from one location to another." In summary, social mobility analysis as we have defined it here is concerned with distinguishing between observed differences with respect

[34]This definition is identical to that offered by Norman Ryder in his unpublished writings. As noted by Fosse (2023), Ryder viewed social structure as a "map of locations" in which individuals (and cohorts) are embedded, while a social process, of which social mobility is a type, is "the aggregate version of movements from one location to another" (5).

to the social structure and social mobility, rather than with identifying unique "effects" of origin, destination, and mobility.

It is crucial to understand that social mobility analysis is distinct from any analysis using models of the form $Y = f(O, D) + \epsilon$ or $Y = f(D, M) + \epsilon$. The reason is that conditioning on class destination conflates variations in social structure with variations in social mobility. For example, suppose we attempt to describe a mobility table using a $f(O, D)$ function, such as Duncan's SAM (without assuming that the estimates represent unique causal effects). Conditional on class destination (D), variations in class origin (O) will compare different origin groups with distinctly different mobility levels, producing a descriptive pattern that captures neither differences in social structure nor social mobility, but a heterogeneous mixture of the two.[35] It is only in the complete absence of social mobility that variations in class origin conditional on class destination will provide meaningful estimates of sociostructural differences.[36]

Likewise, suppose we attempt to specify a model based on a function $f(D, M)$. Conditional on class destination (D), variations across mobility levels (M) will compare groups with distinctly different class origins and mobility levels. Importantly, these estimates will not reflect a social process that any group of individuals will ever experience, and instead conflates sociostructural differences with the social mobility gradient.[37] Similar to the origin-destination model, the destination-mobility model will give meaningful estimates of variability with respect to social mobility only in the absence of sociostructural differences.[38]

Our approach is also distinct from one-factor models of the form $f(O)$, $f(D)$, and $f(M)$. As we show in Online Appendix C, each of these models can be understood as reflecting underlying differences in the social structure as well as social mobility. For example, a model with just class destination is essentially differences in social structure and social mobility, but weighted by the associations between class destination with class origin and mobility, respectively. Among these three general class of models, as noted above, our preference is for models that include just class destination. The reason is that models with just class origin or class mobility are a weighted combination

[35]Note, however, that variations in class destination conditional on class origin will generally produce meaningful information about social mobility, although indexed by class destination rather than mobility levels.

[36]Note that this assumption can be tested against the data (see Online Appendix C for details).

[37]Note again that variations in class destination conditional on class mobility will provide meaningful information about social structure, but indexed by class destination rather than class origin.

[38]This, again, is an assumption that can be tested against the data.

of variability with respect to either social structure or social mobility, along with a set of intra-destination differences. Again, these models conflate sociostructural differences with processes of social mobility. An additional advantage of models of the form $f(D)$ is that there is a simple way to decompose the aggregate class destination gap into a part attributable to the social structure and a part attributable to social mobility, as we outline below.

The injunction to use functions of the form $f(O, M)$ and $f(D)$ to examine the structural and dynamic inequalities in a mobility table is quite general, and in principle any number of modeling approaches could be used. In the next section we introduce a particularly useful model for conducting a social mobility analysis. As we show, this model parsimoniously provides a number of highly informative summaries about the nature and extent of observed differences with respect to both the social structure and social mobility. Moreover, this model extends quite straightforwardly to more complex settings, such as those that involve multiple time points and/or continuous measures.

2. The Structural and Dynamic Inequality Model

To examine differences with respect to social structure and social mobility, we introduce what we call the *Structural and Dynamic Inequality Model*, or, for short, the SDI model. This model is derived by taking the L-ODM model, substituting destination with origin and mobility, and rearranging the terms. The result is a fully-identified model, expressed in terms of the origin, destination, and mobility parameters, with the following general form:

$$\mu_{rijk} = f(O, M) + \epsilon = \mu + \Gamma_1(i - i^*) + \Gamma_2(k - k^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{[i+k-I]} + \widetilde{\gamma}_k + \eta_{i[i+k-I]k} + \xi_{ri[i+k-I]k}, \quad (15)$$

where the asterisks denote midpoint or referent indices $i^* = (I + 1)/2$, $j^* = (J + 1)/2$, and $k^* = (K + 1)/2$; where $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$; $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\gamma}$ represent the origin, destination, and mobility nonlinearities, respectively; $\eta_{i[i+k-I]k}$ is an additional (orthogonal) term denoting interactions; and $\xi_{ri[i+k-I]k}$ is a normally distributed individual-level error term with a mean of zero. Note that, by substituting the destination indices with the sum of the origin and mobility indices (that is, $j = i+k-I$ and $J = K-I+1$), the outcome is only a function of class origin, indexed by i, with corresponding parameters representing differences with respect to the social structure, and mobility, indexed by k, with corresponding parameters representing differences with respect to social mobility. Crucially, this model is identified (i.e., the design matrix is of full rank) as it does not contain a separate linear term for destination, which is instead absorbed into linear terms indexed by origin and mobility.[39]

Somewhat remarkably, Equation 15 encompasses a great deal of information about the underlying structural and dynamic inequalities in a mobility table. Accordingly, a large number of parametric expressions can be derived from the model. Among the most useful for describing the main patterns in a mobility table are those discussed and visualized in the remaining sections: the *Social Structure Matrix* and *Social Mobility Matrix*, the *Social Structure Slope and the Social Mobility Slope*, *Social Structure and Social Mobility Curves*, and *Comparative Mobility Curves*. In Online Appendix G, we include a full list of expressions that can be derived from the SDI model (Table G.1) and illustrate the analytic use of one additional expression (*Marginal Class Destination Curve*).

3. The Social Structure and Mobility Matrices

The SDI model allows us to decompose the overall pattern observed in a mobility table into two distinct underlying matrix-based components. This decomposition helps separate inequalities associated with social positions from those associated with movement between these positions. First, we can extract the *social structure matrix*. This matrix reflects the joint origin-destination parameters from the SDI model and is calculated as: $\widehat{\mu}_{ijk} = \mu + \Gamma_1(i - i^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{[i+k-I]}$. This matrix isolates the component of inequality related to the combination of starting position (origin) and ending position (destination), effectively representing the patterns associated purely with the social structure itself. Second, we can extract the *social mobility matrix*. This matrix reflects the joint mobility-destination parameters from the SDI model, calculated as: $\widehat{\mu}_{ijk} = \mu + \Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{[i+k-I]}$ This matrix captures the component of inequality related to the experience of mobility (the movement itself) in conjunction with the destination reached, representing the dynamic aspect of inequality.

[39] The SDI model can be estimated using a design matrix similar to that of the L-ODM model, but with the destination linear component dropped. Several characteristics of its structure are pertinent: (1) the model inherently incorporates nonlinearities for origin, destination, and mobility; (2) although additive in its form (see Equation 15), the SDI model, as noted previously, exhibits an "interactive" characteristic in practice, as destination nonlinearities, for example, manifest at different mobility levels for different origin groups, thereby functioning as structured origin-mobility interactions; and (3) the cell-specific error term can be interpreted as capturing additional higher-level interactions orthogonal to the main specified terms, the modeling of which is further detailed in Online Appendix B. More generally, the SDI framework can also be modified to model interactions more explicitly. With categorical data, for example, destination nonlinearities can be replaced by specified origin-mobility interactions using orthogonal polynomials, while with continuous variables, one can model origin-mobility interactions using tensor products of smooth functions.

Figure 7 provides a visualization of these components of a mobility table using Sobel's (1981) fertility data. Panels (a) and (b) display the overall predicted fertility values from the main parameters of the SDI model in 2D and 3D, respectively, showing the combined influence of social structure and social mobility. Panel (c) presents the *social structure matrix*, illustrating how predicted fertility varies according to the joint combination of origin and destination positions, thereby highlighting structural inequalities. Panel (d) displays the *social mobility matrix*, showing how predicted fertility varies based on the joint combination of mobility experience and class destination, highlighting dynamic inequalities. Comparing panels (c) and (d) visually separates the relative contributions of structural and dynamic inequalities to the fertility patterns. The social structure matrix (panel c) exhibits marked gradients, especially the trend towards lower fertility with higher origin-destination status, suggesting a prominent role for structural position. Similarly, the social mobility matrix (panel d) illustrates significant fertility differences across various origin-mobility combinations, indicating that the dynamic component also accounts for a considerable part of the overall observed variation.

4. The Social Structure Slope and the Social Mobility Slope

At a more fundamental level, the two linear parameters of interest in Equation 15 are Γ_1 and Γ_2.[40] The linear term Γ_1 is the *social structure slope*, or for short, the *ST slope*, which describes a set of overall (linear) differences in the social structure, or a structural inequality. Note that Γ_1 is the sum of two terms from the L-ODM model: α, the linear origin term, and β, the linear destination term. This is because, within a given mobility level, as we compare successive class origin groups, they are also observed in higher class destination groups. By contrast, Γ_2 is the *social mobility slope*, or, for short, the *SM slope*, which describes a set of overall (linear) differences with respect to social mobility, or a dynamic inequality. Like the ST slope, Γ_2 is the sum of two terms from the L-ODM model: γ, the linear mobility term, and β, the linear destination term. Again, this is because within a given class origin group, greater movement up (or down) the class hierarchy corresponds to a higher (or lower) class destination. With respect to Sobel's (1981) fertility data, the ST slope is $\widehat{\Gamma}_1 =$

[40]Note that, unlike in our discussion of bounding analyses, our goal is to interpret these parameters directly as descriptive quantities, rather than use them to construct a canonical solution line and bounding formulas.

-0.317, while the SM slope is $\widehat{\Gamma}_2 = -0.213$.[41] Both slopes are negative, indicating that higher levels in the social structure as well as upward social mobility both correspond to lower levels of fertility (see also Figure A.2 in Online Appendix A).

The ST and SM slopes can take on different values, with different implications for the patterns observed in a mobility table (see Table A.3 in Online Appendix A). In a world with no social mobility (i.e., where the SM slope is zero), linear differences within different class destinations are the same as linear differences within different mobility groups (i.e., both will reflect the ST slope). Visually, this will appear as a set of linear differences that vary only by class origin, or across the rows of an origin-destination mobility table. Conversely, in a world with no structural inequality (i.e., the ST slope is zero), the linear differences within class origin groups will be the same as those within class destination groups (i.e., both will reflect the SM slope). Visually, this will appear as a set of linear differences that vary across mobility groups, or across the diagonals of an origin-destination mobility table.[42] To our knowledge, this feature of mobility tables has not been recognized before.

5. Social Structure and Social Mobility Curves

The ST and SM slopes provide basic overall information on the overall structural inequality and social mobility gradient observed in the data. However, more informative summaries incorporate the origin and mobility nonlinearities, respectively. Specifically, adding the origin nonlinearities along with the ST slope results in the *social structure curve*, or ST curve, defined by $\Gamma_1(i - i^*) + \widetilde{\alpha}_i$ for origin groups $i = 1, \ldots, I$.[43] Likewise, adding the mobility nonlinearities along with the SM slope results in the *social mobility curve*, or SM curve, defined by $\Gamma_2(k - k^*) + \widetilde{\gamma}_k$ for mobility groups $k = 1, \ldots, K$. The ST and SM curves for Sobel's fertility data are shown in Figure 8. Regarding structural inequality, the main conclusion is similar to that for the ST slope, with higher (or lower) positions corresponding to lower fertility levels, although with few differences between the top two

[41] Note that previously these parameters were discussed only because they define the location of the canonical solution line in the parameter space. Here our focus is on their substantive meaning as descriptive quantities in their own right.

[42] This relationship again mean that one should exercise caution when trying to interpret any observed pattern in a mobility table as purely an origin "effect" or a mobility "effect," because such patterns can appear naturally in the data even though they are joint sets of parameters. It also means that we can plot the ST and SM slopes with respect to class origin and destination, respectively, without loss of information.

[43] For simplicity of presentation, we exclude the intercept from the definitions of the various parametric expressions in the following sections. However, visualizations of these expressions include the intercept term.

social classes. Likewise, we find that downward (or upward) mobility corresponds with lower (or higher) levels of fertility.

6. Comparative Mobility Curves

The ST and SM curves (as well as the slopes) are highly informative, compact summaries of the variability on a mobility table. However, in practice it is often useful to supplement these measures with mobility curves specific to each class origin group, or what we call *comparative mobility curves*. Two particularly useful versions of these curves are shown in Figure 9, again using the fertility data. First, there are what we call *overall comparative mobility curves*, which are defined as:

$$\overbrace{\Gamma_1(i-i^*)+\widetilde{\alpha}_i}^{\substack{\text{Structural Inequality}\\ \text{(Social Structure)}}} + \underbrace{\Gamma_2(k-k^*)+\widetilde{\gamma}_k}_{\substack{\text{Dynamic Inequality}\\ \text{(Social Mobility)}}} \text{ for } k=1,\ldots,K \text{ in each class origin group } i, \tag{16}$$

which represents the pattern of social mobility for a given class origin group i in terms of the *SM curve* (indexed by mobility levels), or $\Gamma_2(k-k^*)+\widetilde{\gamma}_k$, and the overall structural level for a given class origin group i in terms of the *ST curve* (indexed by class origin), or $\Gamma_1(i-i^*)+\widetilde{\alpha}_i$.

A slightly more complex representation is identical to Equation 16 but includes the nonlinearities attributable to class destination. This results in what we call *adjusted comparative mobility curves*:[44]

$$\overbrace{\Gamma_1(i-i^*)+\widetilde{\alpha}_i + \widetilde{\beta}_{i+k-I}}^{\substack{\text{Structural Inequality}\\ \text{(Social Structure)}}} \hspace{-4em} \underbrace{\hphantom{\widetilde{\beta}_{i+k-I}} + \Gamma_2(k-k^*)+\widetilde{\gamma}_k}_{\substack{\text{Dynamic Inequality}\\ \text{(Social Mobility)}}} \text{ for } k=1,\ldots,K \text{ in each class origin group } i, \tag{17}$$

where the $\widetilde{\beta}_{i+k-I}$ terms are the class destination nonlinearities. Note that the destination nonlinearities are indexed by both class origin and mobility levels, and thus represent both structural and dynamic inequalities.

[44]These curves are "adjusted" in the specific sense that they are purged of cell-specific heterogeneity (i.e., joint interaction terms). In general, the raw (or "unadjusted") cell means (or other summaries) of a mobility table will differ from the "adjusted" means given by Equation 17. Note that, to keep our analyses consistent with our analyses of the DRM, we are using simulated data in which the proportionality constraint of the DRM holds with respect to the nonlinearities. Thus, in our analyses below the "adjusted" curves are equal to the "unadjusted" curves (see Table G.1 in Online Appendix G).

Figure 9 shows overall and adjusted comparative mobility curves using Sobel's (1981) fertility data. Panel (a) of Figure 9 displays the collection of overall comparative mobility curves (Equation 16). Each of the class origin groups are labeled from 1 to 5. The vertical axis is the number of children, while the x-axis is the mobility dimension. Note that each class origin group is entitled to only a specific pattern of social mobility. For example, those at the top of the class hierarchy (class 5) can either be immobile or move downward, while those at the bottom of the class hierarchy (class 1) can either be immobile or move upward.

These curves encapsulate a great deal of information about the relationships between the social structure and social mobility with respect to fertility. Vertical differences between the class origin groups are a function of the structural component, or $\Gamma_1(i - i^*) + \widetilde{\alpha}_i$. For a given origin group i, these terms effectively act like an intercept that moves the mobility curves up and down. If social structure were irrelevant for fertility, then all of the mobility curves would collapse on top of each other, and thus a single mobility curve would adequately describe all of the mobility patterns. Likewise, panel (a) of Figure 8, which shows the ST curve, would be a flat line. Horizontal differences in a given class origin group are a function of the social mobility component, or $\Gamma_2(k - k^*) + \widetilde{\gamma}_k$. If social mobility were irrelevant for fertility, then each mobility curve would be a horizontal straight line, and all differences would be described in terms of the social structure. Accordingly, panel (b) of Figure 8, which shows the SM curve, would be a flat line.

Panel (b) shows the adjusted comparative mobility curves, which are identical to the overall curves in panel (a) except that the class destination nonlinearities are included. Vertical and horizontal differences in the graph still represent structural and dynamic inequalities, respectively, but they are now represented in a more complex way. Specifically, vertical differences between the class origin groups are now a function of $\Gamma_1(i - i^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{i+k-I}$, while horizontal differences in a given class origin group are a function $\Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{i+k-I}$. Note that the class destination nonlinearities, although additive in the model, are "interactive" in the data in that they appear at different mobility levels for different class origin groups (for the overall destination nonlinearities, see Figure A.1 (b) in Online Appendix A). This is a fundamental feature of data from a mobility table that, to our knowledge, has not been previously acknowledged. What this means in practice is that a model that is additive in the parameters, such as the SDI model, is, in fact, "interactive" in the data.[45]

Two fundamental conclusions are apparent from the mobility curves shown in Figure 9. First, conditional on social mobility, differences in the social structure are systematically related to lower levels of fertility, particularly between classes 1 and 2 as well as 3 and 4. Second, conditional on the social structure, upward (or downward) social mobility is systematically related to lower (or higher) levels of fertility for all class origin groups. Note that these conclusions hold even for the adjusted mobility curves shown in panel (b), which injects additional heterogeneity into the patterns.

The results in this section underscore that there is much to be learned from a purely descriptive analysis of a simple mobility table. Rather than focusing on using information external to the data to extract unique "effects" for origin, destination, and mobility, as is common in the existing literature on mobility effects, our focus has been on estimating quantities that represent various aspects of structural and dynamic inequalities. This has the great advantage of facilitating the accumulation of knowledge, as the underlying models are identified.

Conclusion

Despite decades of research, quantitative evidence on the consequences of social mobility has been inconclusive as it has been hampered by longstanding methodological challenges. In this article, we make a number of important contributions to the literature on social mobility effects, which has been dominated in recent years by the Diagonal Reference Model (DRM).

First, we show that for plausible values of mobility effects, the DRM will in general implicitly force the underlying linear mobility effect to zero. In addition, we show both mathematically and through simulations that the mobility effects estimated by the DRM are sensitive to the size and sign of the origin and destination linear effects, often in ways that may be counterintuitive to applied researchers. This finding may account for the fact that applied researchers have generally found weak or no evidence of mobility effects on a wide range of outcomes. Second, we generalize the identification problem of conventional mobility effect models by showing that the DRM and related methods can be viewed as special cases of a bounding analysis, where point identification is achieved by invoking extremely strong assumptions (resulting in very tight bounds).

[45] Note, however, that while the class destination nonlinearities appear at different mobility levels for different class origin groups, magnitude and direction of the class destinations are the same.

Finally, we present a new framework for the analysis of mobility tables based on the identification and estimation of joint parameter sets, introducing what we call the *Structural and Dynamic Inequality* (SDI) model. This model is fully identified, relies on much weaker assumptions than conventional models of mobility effects, and can be treated both as a descriptive model and, if additional assumptions are invoked, as a causal model. We propose a range of new, highly informative, and easily interpretable estimates and graphical approaches that compare structural and dynamic aspects of inequality. These estimates and graphical approaches show that there is much to be learned from a purely descriptive analysis of a simple mobility table and should serve as a useful tool for future research on social mobility.

An important and perhaps controversial conclusion from this article is that researchers should avoid using models that attempt to extract unique omnibus "effects" for origin, destination, and mobility. The assumptions behind these models are not only very strong but also not directly testable against the data. Of course, assumptions are fundamental to all statistical modeling, including standard approaches like ordinary least squares regression. However, we have sought to show that the necessary assumptions to identify mobility effects are much stronger than generally recognized. As a result, we consider the evidentiary basis for the individual-level effects of social mobility highly debatable. Not because this rapidly expanding literature, which overwhelmingly uses the DRM, has produced mixed and often null findings. But because using a model that is unidentified carries with it very strong, untestable assumptions, with parameters that are compatible with an infinite number of solutions. If researchers still wish to identify omnibus factors for origin, destination, and mobility, we would recommend using bounding analyses that make *explicit* assumptions about the size, sign, and/or shape of the effects. Ideally, these assumptions should be driven by social theory or background knowledge, and the conclusions should be stated as inherently tentative.

More importantly, however, we call for a fundamental shift in the analytic focus of mobility research. This shift rests on the basic insights that it is not meaningful to divorce class destination from class mobility. We consider the most appropriate goal of social mobility analysis the identification of joint sets of parameters, such as the joint effect of social destination and social mobility. We hope that this shift in focus will move the literature to more meaningful answers that can be falsified against the data. Our new approach, consistent with what we term "positional sociology," is to treat origin, destination, and mobility as dimensions along which variability is observed, rather than as proxies for unobserved causal factors (Fosse and Pfeffer 2023). Interpreting origin, destination, and mobility as dimensions of variability, and working with functions of the form $f(O, M)$, are the core ideas of the kind of social mobility analysis we envision.

At the same time, we acknowledge that, despite the aforementioned discussion, some analysts will persist in attempting to isolate distinct "effects" for origin, destination, and mobility. To help guide applied researchers, we outline a sequence of steps for analyzing data from a social mobility table, whether one's aim is to describe structural and dynamic inequalities based on identifiable, joint parameter sets or to extract unique effects for origin, destination, and mobility. Table 7 summarizes these steps and calls out the sections or equations in the manuscript that match each step.

First, researchers should fit the Structural and Dynamic Inequality (SDI) model, as detailed in Section III.2 and Equation 15. This model serves as the foundation for understanding the distinct components of inequality. Second, the interpretation phase involves examining the Social Structure (ST) and Social Mobility (SM) slopes, the ST and SM curves, as well as the Social Structure and Social Mobility matrices. These elements, discussed in Sections III.4 and III.5, provide fundamental insights into the structural and dynamic aspects of inequality. Third, a more detailed examination involves interpreting the overall and adjusted comparative mobility curves, as explained in Section III.6 and outlined in Equations 16 and 17. These curves illustrate specific mobility patterns for different origin groups. Fourth, if the goal is to identify the unique "effects" of origin, destination, and mobility, then we recommend that researchers adopt a partial identification approach in which the parameters are bounded based on explicit assumptions about the size, sign, or shape of the effects. This method is covered in Sections II.1 and II.2. Finally, when employing bounding strategies, it is crucial to provide a range of these strategies and to emphasize the inherently tentative, non-falsifiable nature of the conclusions. Moreover, the strong assumptions required to interpret these quantities in terms of formal counterfactuals should be clearly articulated, as discussed in Section II.2 and Online Appendix F.

We end by outlining a few options for future research that could extend the SDI model in a variety of different directions: First, the SDI model can be extended by interacting the components with time, allowing one to examine how the roles of social structure and social mobility change over time. Note, however, that this requires a large sample due to the sparseness at the extremes of the mobility table, as well as appropriate restrictions to deal with overfitting noise. Second, another extension is to include covariates in the model and examine potential interactions. Third, researchers should consider extending the SDI model to continuous data. This will require models that use more flexible functional forms, such as cubic regression splines or generalized additive models (GAMs), to deal with complex nonlinearities in the data. Fourth, the SDI model can be extended to address directional mobility, i.e., to distinguish the distinct patterns associated with downward and upward mobility. As currently formulated, the model captures the average linear component representing mobility-related differences (via the SM slope, Γ_2), while accommodating asymmetric overall patterns of mobility through its flexible nonlinear mobility components. One could conceive of extensions where this linear component itself differs systematically by direction (upward versus downward mobility). Potential modifications could involve specifying separate linear SM slopes for upward and downward mobility, possibly through interaction terms with a directional indicator or by using explicit piecewise coding. Finally, more generally, we recommend that researchers consider alternative ways of specifying models of the form $Y = f(O, M) + \epsilon$ (as well as $Y = f(D) + \epsilon$). Although we believe that the SDI model is particularly useful for summarizing data from a mobility table, other possible models, such as those that explicitly incorporate interactions between class origin and mobility, may be useful in other applications.

References

Bijlsma, Maarten J., Rhian M. Daniel, Fanny Janssen, and Bianca L. De Stavola. 2017. "An Assessment and Extension of the Mechanism-Based Approach to the Identification of Age-Period-Cohort Models." *Demography* 54:721–743. ISSN 0070-3370, 1533-7790. doi:10.1007/s13524-017-0562-6.

Billingsley, Sunnee and Anna Matysiak. 2018. "Social Mobility and Family Expansion in Poland and Russia during Socialism and Capitalism." *Advances in Life Course Research* 36:80–91. ISSN 1040-2608. doi:10.1016/j.alcr.2018.04.001.

Blalock, H. M. 1966. "Comment: Status Inconsistency and the Identification Problem." *Public Opinion Quarterly* 30:130. ISSN 0033362X. doi:10.1086/267387.

———. 1967. "Status Inconsistency, Social Mobility, Status Integration and Structural Effects." *American Sociological Review* 32:790–801. ISSN 00031224. doi:10.2307/2092026.

Bloome, Deirdre. 2015. "Income Inequality and Intergenerational Income Mobility in the United States." *Social Forces* 93:1047–1080. ISSN 0037-7732, 1534-7605. doi:10.1093/sf/sou092.

Breen, Richard. 2001. "Social Mobility and Constitutional and Political Preferences in Northern Ireland." *The British Journal of Sociology* 52:621–645. ISSN 1468-4446. doi:10.1080/00071310120084508.

Breen, Richard and Christopher T. Whelan. 1994. "Social Class, Class Origins and Political Partisanship in the Republic of Ireland." *European Journal of Political Research* 26:117–133. ISSN 1475-6765. doi:10.1111/j.1475-6765.1994.tb00436.x.

Bulczak, Grzegorz and Alexi Gugushvili. 2022. "Downward Income Mobility among Individuals with Poor Initial Health Is Linked with Higher Cardiometabolic Risk." *PNAS Nexus* 1:pgac012. ISSN 2752-6542. doi:10.1093/pnasnexus/pgac012.

Bulczak, Grzegorz, Alexi Gugushvili, and Olga Zelinska. 2022. "How Are Social Origin, Destination and Mobility Linked to Physical, Mental, and Self-Rated Health? Evidence from the United States." *Quality & Quantity* 56:3555–3585. ISSN 1573-7845. doi:10.1007/s11135-021-01286-5.

Chan, Tak Wing and Heather Turner. 2017. "Where Do Cultural Omnivores Come from? The Implications of Educational Mobility for Cultural Consumption." *European Sociological Review* 33:576–589. ISSN 0266-7215. doi:10.1093/esr/jcx060.

Chen, Edith, Gene H. Brody, and Gregory E. Miller. 2022. "What Are the Health Consequences of Upward Mobility?" *Annual Review of Psychology* 73:599–628. doi:10.1146/annurev-psych-033020-122814.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129:1553–1623. ISSN 0033-5533, 1531-4650. doi:10.1093/qje/qju022.

Clifford, P. and A. F. Heath. 1993. "The Political Consequences of Social Mobility." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 156:51–61. ISSN 0964-1998. doi:10.2307/2982860.

Clogg, Clifford C. 1982. "Cohort Analysis of Recent Trends in Labor Force Participation." *Demography* 19:459–479.

Coulangeon, Philippe. 2013. "The Omnivore and the 'Class Defector'. Musical Taste and Social Mobility in Contemporary France."

———. 2015. "Social Mobility and Musical Tastes: A Reappraisal of the Social Meaning of Taste Eclecticism." *Poetics* 51:54–68. ISSN 0304-422X. doi:10.1016/j.poetic.2015.05.002.

Cox, D. R. 1990. "Role of Models in Statistical Analysis." *Statistical science* 5:169–174. ISSN 0883-4237. doi:10.1214/ss/1177012165.

Creighton, Mathew J., Daniel Capistrano, and Monika da Silva Pedroso. 2022. "Educational Mobility and Attitudes Towards Migration from an International Comparative Perspective." *Journal of International Migration and Integration* ISSN 1874-6365. doi:10.1007/s12134-022-00977-8.

Daenekindt, Stijn. 2017. "The Experience of Social Mobility: Social Isolation, Utilitarian Individualism, and Social Disorientation." *Social Indicators Research* 133:15–30. ISSN 0303-8300, 1573-0921. doi:10.1007/s11205-016-1369-3.

Daenekindt, Stijn and Henk Roose. 2013. "A Mise-en-scène of the Shattered Habitus: The Effect of Social Mobility on Aesthetic Dispositions Towards Films." *European Sociological Review* 29:48–59. ISSN 0266-7215. doi:10.1093/esr/jcr038.

———. 2014. "Social Mobility and Cultural Dissonance." *Poetics* 42:82–97. ISSN 0304-422X. doi:10.1016/j.poetic.2013.11.002.

Daenekindt, Stijn, Jeroen van der Waal, and Willem de Koster. 2018. "Social Mobility and Political Distrust: Cults of Gratitude and Resentment?" *Acta Politica* 53:269–282. ISSN 1741-1416. doi:10.1057/s41269-017-0050-4.

De Graaf, Nan Dirk, Paul Nieuwbeerta, and Anthony Heath. 1995. "Class Mobility and Political Preferences: Individual and Contextual Effects." *American Journal of Sociology* 100:997–1027. ISSN 0002-9602. doi:10.1086/230607.

De Graaf, Nan Dirk and Wout Ultee. 1990. "Individual Preferences, Social Mobility and Electoral Outcomes." *Electoral Studies* 9:109–132. ISSN 0261-3794. doi:10.1016/0261-3794(90)90003-Q.

Dennison, Christopher R. 2018. "Intergenerational Mobility and Changes in Drug Use Across the Life Course." *Journal of Drug Issues* 48:205–225. ISSN 0022-0426. doi:10.1177/0022042617746974.

Dhoore, Jasper, Stijn Daenekindt, and Henk Roose. 2019. "Social Mobility and Life Satisfaction across European Countries: A Compositional Perspective on Dissociative Consequences of Social Mobility." *Social Indicators Research* 144:1257–1272. ISSN 1573-0921. doi:10.1007/s11205-019-02083-2.

Domański, Henryk and Zbigniew Karpiński. 2018. "Intergenerational Mobility and Omnivorism in Eating." *Appetite* 121:83–92. ISSN 0195-6663. doi:10.1016/j.appet.2017.10.030.

Duncan, Otis Dudley. 1966. "Methodological Issues in the Study of Social Sobility." In *Social Structure and Mobility in Economic Development*, pp. 51–97. Chicago: Aldine.

Engzell, Per and Mathieu Ichou. 2020. "Status Loss: The Burden of Positively Selected Immigrants." *International Migration Review* 54:471–495. ISSN 0197-9183. doi:10.1177/0197918319850756.

Erikson, Robert and John H Goldthorpe. 1992. *The Constant Flux. A Study of Class Mobility in Industrial Societies*. Oxford: Clarendon Press.

Fan, Xiaoguang and Fei Yan. 2019. "The Long Shadow: Social Mobility and Political Participation in Urban China, 2006–2012." *Social Science Research* ISSN 0049-089X. doi:10.1016/j.ssresearch.2019.03.006.

Fosse, Ethan. 2021. "Bayesian Age-Period-Cohort Models." In *Age, Period, and Cohort Effects: The Identification Problem and Beyond*, edited by Andrew Bell, pp. 142–175. London, UK: Routledge Press.

———. 2023. "Theorizing in the Cohort Mode: On Ryder's Processual Account of Social Change." *Canadian Studies in Population* 50:5. ISSN 1927-629X. doi:10.1007/s42650-023-00075-9.

Fosse, Ethan and Fabian Pfeffer. 2023. "Manifesto for Positional Sociology." *Data Sciences Institute Working Paper, University of Toronto* .

Fosse, Ethan and Christopher Winship. 2018. "Moore–Penrose Estimators of Age–Period–Cohort Effects: Their Interrelationship and Properties." *Sociological Science* 5:304–334.

———. 2019a. "Analyzing Age-Period-Cohort Data: A Review and Critique." *Annual Review of Sociology* 45:467–492. ISSN 0360-0572, 1545-2115. doi:10.1146/annurev-soc-073018-022616.

———. 2019b. "Bounding Analyses of Age-Period-Cohort Effects." *Demography* 56:1975–2004.

Friedman, Sam. 2016. "Habitus Clivé and the Emotional Imprint of Social Mobility." *The Sociological Review* 64:129–147. ISSN 0038-0261. doi:10.1111/1467-954X.12280.

Goldthorpe, John. 1999. "Modelling the Pattern of Class Voting in British Elections, 1964-1992." In *The End of Class Politics? Class Voting in Comparative Context*, edited by Geoffrey Evans, pp. 59–82. London, UK: Oxford University Press.

Gowen, Ohjae, Ethan Fosse, and Christopher Winship. 2023. "Cross-Group Differences in Age, Period, and Cohort Effects: A Bounding Approach to the Gender Wage Gap." *Sociological Science* 10:731–768. ISSN 2330-6696. doi:10.15195/v10.a26.

Graaf, Nan Dirk De, Paul Nieuwbeerta, and Anthony Heath. 1995. "Class Mobility and Political Preferences: Individual and Contextual Effects." *American Journal of Sociology* ISSN 0002-9602. doi:10.1086/230607.

Graf, Gloria Huei-Jong, Yalu Zhang, Benjamin W Domingue, Kathleen Mullan Harris, Meeraj Kothari, Dayoon Kwon, Peter Muennig, and Daniel W Belsky. 2022. "Social Mobility and Biological Aging among Older Adults in the United States." *PNAS Nexus* 1:pgac029. ISSN 2752-6542. doi:10.1093/pnasnexus/pgac029.

Grusky, David B and Robert M Hauser. 1984. "Comparative Social Mobility Revisited. Models of Convergence and Divergence in 16 Countries." *American Sociological Review* 49:19–38.

Gugushvili, A., Y. Zhao, and E. Bukodi. 2020. "Intergenerational Educational Mobility and Smoking: A Study of 20 European Countries Using Diagonal Reference Models." *Public Health* 181:94–101. ISSN 0033-3506. doi:10.1016/j.puhe.2019.12.009.

Gugushvili, Alexi, Grzegorz Bulczak, Olga Zelinska, and Jonathan Koltai. 2021. "Socioeconomic Position, Social Mobility, and Health Selection Effects on Allostatic Load in the United States." *PLOS ONE* 16:e0254414. ISSN 1932-6203. doi:10.1371/journal.pone.0254414.

Gugushvili, Alexi, Yizhang Zhao, and Erzsébet Bukodi. 2019. "'Falling from Grace' and 'Rising from Rags': Intergenerational Educational Mobility and Depressive Symptoms." *Social Science & Medicine* 222:294–304. ISSN 0277-9536. doi:10.1016/j.socscimed.2018.12.027.

Heckman, James and Richard Robb. 1985. "Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations." In *Cohort Analysis in Social Research*, edited by Stephen E. Fienberg and William M. Mason, pp. 137–150. New York, NY: Springer.

Hendrickx, John, Nan Dirk De Graaf, Jan Lammers, and Wout Ultee. 1993. "Models for Status Inconsistency and Mobility: A Comparison of the Approaches by Hope and Sobel with the Mainstream Square Additive Model." *Quality and Quantity* 27:335–352. ISSN 0033-5177, 1573-7845. doi:10.1007/BF01102497.

Hernán, Miguel A. 2016. "Does Water Kill? A Call for Less Casual Causal Inferences." *Annals of Epidemiology* 26:674–680. ISSN 10472797. doi:10.1016/j.annepidem.2016.08.016.

Hope, Keith. 1971. "Social Mobility and Fertility." *American Sociological Review* 36:1019–1032. ISSN 0003-1224. doi:10.2307/2093762.

———. 1975. "Models of Status Inconsistency and Social Mobility Effects." *American Sociological Review* 40:322–343. ISSN 0003-1224. doi:10.2307/2094461.

Houle, Jason N. and Molly A. Martin. 2011. "Does Intergenerational Mobility Shape Psychological Distress? Sorokin Revisited." *Research in Social Stratification and Mobility* 29:193–203. ISSN 0276-5624. doi:10.1016/j.rssm.2010.11.001.

Iveson, Matthew H., Simon R. Cox, and Ian J. Deary. 2022. "Intergenerational Social Mobility and Health in Later Life: Diagonal Reference Models Applied to the Lothian Birth Cohort 1936." *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* p. gbac107. ISSN 1758-5368. doi:10.1093/geronb/gbac107.

Jackson, John W. and Tyler J. VanderWeele. 2018. "Decomposition Analysis to Identify Intervention Targets for Reducing Disparities." *Epidemiology* 29:825–835.

Jaime-Castillo, Antonio M. and Ildefonso Marqués-Perales. 2019. "Social Mobility and Demand for Redistribution in Europe: A Comparative Analysis." *The British Journal of Sociology* 70:138–165. ISSN 1468-4446. doi:10.1111/1468-4446.12363.

Kaiser, Caspar and Nhat An Trinh. 2021. "Positional, Mobility, and Reference Effects: How Does Social Class Affect Life Satisfaction in Europe?" *European Sociological Review* ISSN 0266-7215. doi:10.1093/esr/jcaa067.

Kempel, Mia Klinkvort, Trine Nøhr Winding, Morten Böttcher, and Johan Hviid Andersen. 2022. "Evaluating the Association between Socioeconomic Position and Cardiometabolic Risk Markers in Young Adulthood by Different Life Course Models." *BMC Public Health* 22:694. ISSN 1471-2458. doi:10.1186/s12889-022-13158-0.

Kraus, Lisa-Marie and Stijn Daenekindt. 2022. "Moving into Multiculturalism. Multicultural Attitudes of Socially Mobile Individuals without a Migration Background." *European Societies* 24:7–28. ISSN 1461-6696. doi:10.1080/14616696.2021.1976415.

Kulis, Stephen. 1987. "Socially Mobile Daughters and Sons of the Elderly: Mobility Effects within the Family Revisited." *Journal of Marriage and Family* 49:421–433. ISSN 0022-2445. doi:10.2307/352311.

Kwon, Aram. 2022. "The Impact of Intergenerational Mobility on Well-being in Japan." *Social Indicators Research* 162:253–277. ISSN 1573-0921. doi:10.1007/s11205-021-02834-0.

Lipset, Seymour Martin and Reinhard Bendix. 1959. *Social Mobility in Industrial Society*. Transaction Publishers. ISBN 978-1-4128-3435-3.

Lipset, Seymour Martin and Hans L Zetterberg. 1959. "Social Mobility in Industrial Society." In *Social Mobility in Industrial Society*, edited by Seymour Martin Lipset and Reinhard Bendix, pp. 11–75. Berkeley: University of California Press.

Luo, Liying. 2022. "Heterogeneous Effects of Intergenerational Social Mobility: An Improved Method and New Evidence." *American Sociological Review* 87:143–173. ISSN 0003-1224. doi:10.1177/00031224211052028.

Manski, Charles F. 1990. "Nonparametric Bounds on Treatment Effects." *The American Economic Review* 80:319–323. ISSN 0002-8282.

———. 2003. "Identification Problems in the Social Sciences and Everyday Life." *Southern Economic Journal* 70:11–21. ISSN 0038-4038. doi:10.2307/1061629.

Marshal, Gordon and David Firth. 1999. "Social Mobility and Personal Satisfaction: Evidence from Ten Countries." *The British Journal of Sociology* 50:28–48.

McNeil, Andrew. 2022. "Intergenerational Social Mobility and Anti-System Support: The Journey Matters." https://www.lse.ac.uk/International-Inequalities/Publications/All-LSE-III-Working-Papers.

Mijs, Jonathan J.B., Stijn Daenekindt, Willem de Koster, and Jeroen van der Waal. 2022. "Belief in Meritocracy Reexamined: Scrutinizing the Role of Subjective Social Mobility." *Social Psychology Quarterly* p. 01902725211063818. ISSN 0190-2725. doi:10.1177/01902725211063818.

Missinne, Sarah, Stijn Daenekindt, and Piet Bracke. 2015. "The Social Gradient in Preventive Healthcare Use: What Can We Learn from Socially Mobile Individuals?" *Sociology of Health & Illness* 37:823–838. ISSN 1467-9566. doi:10.1111/1467-9566.12225.

Monden, Christiaan W.S. and Nan Dirk de Graaf. 2013. "The Importance of Father's and Own Education for Self-Assessed Health across Europe: An East–West Divide?" *Sociology of Health & Illness* 35:977–992. ISSN 1467-9566. doi:10.1111/1467-9566.12015.

Nieuwbeerta, Paul. 1995. *The Democratic Class Struggle in Twenty Countries, 1945-1990.* Thesis Publishers. ISBN 978-90-5170-336-8.

Nieuwbeerta, Paul, Nan Dirk de Graaf, and Wout Ultee. 2000. "The Effects of Class Mobility on Class Voting in Post-War Western Industrialized Countries." *European Sociological Review* 16:327–348. ISSN 0266-7215. doi:10.1093/esr/16.4.327.

Nieuwbeerta, P. and Graaf, N.D. de. 1993. "Intergenerational Class Mobility and Political Preferences in the Netherlands between 1970 and 1986." *Exchanges: The Warwick Research Journal* 29:28–45.

Paterson, Lindsay. 2008. "Political Attitudes, Social Participation and Social Mobility: A Longitudinal Analysis1." *The British Journal of Sociology* 59:413–434. ISSN 1468-4446. doi: 10.1111/j.1468-4446.2008.00201.x.

Pearl, Judea. 2009. *Causality : Models, Reasoning, and Inference / Judea Pearl.*, volume 2nd ed. New York: Cambridge University Press. ISBN 978-0-521-89560-6.

Präg, Patrick and Lindsay Richards. 2018. "Intergenerational Social Mobility and Allostatic Load in Great Britain." *J Epidemiol Community Health* pp. jech–2017–210171. ISSN 0143-005X, 1470-2738. doi:10.1136/jech-2017-210171.

Rehkopf, David H., M. Maria Glymour, and Theresa L. Osypuk. 2016. "The Consistency Assumption for Causal Inference in Social Epidemiology: When a Rose Is Not a Rose." *Current epidemiology reports* 3:63–71. ISSN 2196-2995. doi:10.1007/s40471-016-0069-5.

Robins, James M., Miguel Ángel Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology:." *Epidemiology* 11:550–560. ISSN 1044-3983. doi:10.1097/00001648-200009000-00011.

Rotengruber, Przemysław and Juliusz Tyszka. 2021. "Cultural Course Correction or Back to the Past? | Institute of Cultural Studies of Adam Mickiewicz University in Poznań."

Ryle, Gilbert. 1959. *The Concept of Mind*. London, UK: Hutchinson Press. ISBN 978-0-389-00232-1.

Schaeffer, Merlin. 2019. "Social Mobility and Perceived Discrimination: Adding an Intergenerational Perspective." *European Sociological Review* 35:65–80. ISSN 0266-7215. doi:10.1093/esr/jcy042.

Schuck, Bettina and Nadia Steiber. 2018. "Does Intergenerational Educational Mobility Shape the Well-Being of Young Europeans? Evidence from the European Social Survey." *Social Indicators Research* 139:1237–1255. ISSN 1573-0921. doi:10.1007/s11205-017-1753-7.

Sieben, Inge. 2017. "Child-Rearing Values: The Impact of Intergenerational Class Mobility." *International Sociology* 32:369–390. ISSN 0268-5809. doi:10.1177/0268580917693954.

Smith, Herbert L. 2021. "Age-Period-Cohort Analysis: What Is It Good For?" In *Age, Period and Cohort Effects: Statistical Analysis and the Identification Problem*, pp. 176–205. New York, NY: Routlege.

Sobel, Michael E. 1981. "Diagonal Mobility Models: A Substantively Motivated Class of Designs for the Analysis of Mobility Effects." *American Sociological Review* 46:893–906. ISSN 0003-1224. doi:10.2307/2095086.

———. 1985. "Social Mobility and Fertility Revisited: Some New Models for the Analysis of the Mobility Effects Hypothesis." *American Sociological Review* 50:699–712. ISSN 0003-1224. doi: 10.2307/2095383.

Sorokin, Pitirim A. 1927. *Social Mobility*, volume 4. New York: Harper & Brothers.

Steiber, Nadia. 2019. "Intergenerational Educational Mobility and Health Satisfaction across the Life Course: Does the Long Arm of Childhood Conditions Only Become Visible Later in Life?" *Social Science & Medicine* 242:112603. ISSN 0277-9536. doi:10.1016/j.socscimed.2019.112603.

Swift, Adam. 2004. "Would Perfect Mobility Be Perfect?" *European Sociological Review* 20:1–11.

Tarrence, Jacob. 2018. *Struggling in The Land of Opportunity: Examining Racial Heterogeneity in The Effects of Intergenerational Educational Mobility On Health*. Ph.D. thesis, The Ohio State University.

Tarrence, Jake. 2022. "Is Educational Mobility Harmful for Health?" *Social Science Research* 107:102741. ISSN 0049-089X. doi:10.1016/j.ssresearch.2022.102741.

Tolsma, Jochem, Nan Dirk De Graaf, and Lincoln Quillian. 2009. "Does Intergenerational Social Mobility Affect Antagonistic Attitudes towards Ethnic Minorities?" *The British Journal of Sociology* 60:257–277. ISSN 1468-4446. doi:10.1111/j.1468-4446.2009.01230.x.

Turner, Tom and Lorraine Ryan. 2023. "Shaping Political Orientations: Testing the Effect of Unemployment on Ideological Beliefs and Voting Behaviour." *Irish Political Studies* 38:189–209. ISSN 0790-7184, 1743-9078. doi:10.1080/07907184.2022.2118719.

van der Waal, Jeroen, Stijn Daenekindt, and Willem de Koster. 2017. "Statistical Challenges in Modelling the Health Consequences of Social Mobility: The Need for Diagonal Reference Models." *International Journal of Public Health* 62:1029–1037. ISSN 1661-8564. doi:10.1007/s00038-017-1018-x.

Vansteelandt, Stijn and Rhian M. Daniel. 2017. "Interventional Effects for Mediation Analysis with Multiple Mediators:." *Epidemiology* 28:258–265. ISSN 1044-3983. doi:10.1097/EDE.0000000000000596.

Vansteelandt, Stijn and Marshall Joffe. 2014. "Structural Nested Models and G-estimation: The Partially Realized Promise." *Statistical Science* 29. ISSN 0883-4237. doi:10.1214/14-STS493.

Veenstra, Gerry and Adam Vanzella-Yang. 2021. "Intergenerational Social Mobility and Self-Rated Health in Canada." *SSM - Population Health* 15:100890. ISSN 2352-8273. doi:10.1016/j.ssmph.2021.100890.

Weakliem, David L. 1992. "Does Social Mobility Affect Political Behaviour?" *European Sociological Review* 8:153–165. ISSN 0266-7215. doi:10.1093/oxfordjournals.esr.a036629.

Wiertz, Dingeman and Toni Rodon. 2021. "Frozen or Malleable? Political Ideology in the Face of Job Loss and Unemployment." *Socio-Economic Review* 19:307–331. ISSN 1475-1461. doi:10.1093/ser/mwz024.

Wilson, George, Vincent Roscigno, Carsten Sauer, and Nick Petersen. 2022. "Mobility, Inequality, and Beliefs About Distribution and Redistribution." *Social Forces* 100:1053–1079. ISSN 0037-7732. doi:10.1093/sf/soab047.

Wodtke, Geoffrey T. and Xiang Zhou. 2020. "Effect Decomposition in the Presence of Treatment-induced Confounding: A Regression-with-residuals Approach." *Epidemiology* 31:369–375. ISSN 1044-3983. doi:10.1097/EDE.0000000000001168.

Wright, Erik Olin, ed. 2005. *Approaches to Class Analysis*. Cambridge: Cambridge University Press.

Yang, Xiaozhao. 2016. "China Twenty Years After: Substance Use Under Rapid Social Changes." *Open Access Dissertations* .

Yang, Xiaozhao Yousef. 2020. "Class Status and Social Mobility on Tobacco Smoking in Post-Reform China Between 1991 and 2011." *Nicotine & Tobacco Research* 22:2188–2195. ISSN 1469-994X. doi:10.1093/ntr/ntaa103.

Zang, Emma and Nan Dirk de Graaf. 2016. "Frustrated Achievers or Satisfied Losers? Inter- and Intragenerational Social Mobility and Happiness in China." *Sociological Science* 3:779–800. ISSN 23306696. doi:10.15195/v3.a33.

Zang, Emma, Michael E. Sobel, and Liying Luo. 2023. "The Mobility Effects Hypothesis: Methods and Applications." *Social Science Research* 110:102818. ISSN 0049-089X. doi:10.1016/j.ssresearch.2022.102818.

Zelinska, Olga, Alexi Gugushvili, and Grzegorz Bulczak. 2021. "Social Mobility, Health and Wellbeing in Poland." *Frontiers in Sociology* 6:736249. ISSN 2297-7775. doi:10.3389/fsoc.2021.736249.

Zhao, Yizhang and Yaojun Li. 2019. "Differential Acculturation: A Study of Well-Being Differences in Intergenerational Social Mobility between Rural and Urban China." *Sociology* p. 0038038518818405. ISSN 0038-0385. doi:10.1177/0038038518818405.

Tables

Table 1: Diagonal Reference Model on a 3×3 Mobility Table

	$j=1$	$j=2$	$J=3$	
$i=1$	$\mu_{113} = \mu_{11} + \gamma_3 + \phi_{113}$	$\mu_{124} = w_o\mu_{11} + w_d\mu_{22} + \gamma_4 + \phi_{124}$	$\mu_{135} = w_o\mu_{11} + w_d\mu_{33} + \gamma_5 + \phi_{135}$	$K=5$
$i=2$	$\mu_{212} = w_o\mu_{22} + w_d\mu_{11} + \gamma_2 + \phi_{212}$	$\mu_{223} = \mu_{22} + \gamma_3 + \phi_{223}$	$\mu_{234} = w_o\mu_{22} + w_d\mu_{33} + \gamma_4 + \phi_{234}$	$k=4$
$I=3$	$\mu_{311} = w_o\mu_{33} + w_d\mu_{11} + \gamma_1 + \phi_{311}$	$\mu_{322} = w_o\mu_{33} + w_d\mu_{22} + \gamma_2 + \phi_{322}$	$\mu_{333} = \mu_{33} + \gamma_3 + \phi_{333}$	$k=3$
	$k=1$	$k=2$	$k=3$	

Notes: Number of origin and destination groups is set at $I = 3$ and $J = 3$. The DRM is given generically by $\mu_{ijk} = w_o\mu_{i[j=i]} + w_d\mu_{[i=j]j} + \gamma_k + \phi_{ijk}$.

Table 2: Sobel (1981) Fertility Data

	$j=1$	$j=2$	$j=3$	$j=4$	$J=5$	
$i=1$	3.194	2.850	2.760	2.142	2.252	$K=9$
	(484)	(687)	(462)	(148)	(389)	
$i=2$	3.000	2.423	2.423	1.908	2.180	$k=8$
	(34)	(513)	(345)	(163)	(322)	
$i=3$	1.900	2.349	2.407	2.020	2.113	$k=7$
	(20)	(269)	(307)	(153)	(318)	
$i=4$	2.700	2.225	2.389	1.631	2.005	$k=6$
	(10)	(71)	(54)	(65)	(190)	
$I=5$	2.850	2.308	2.031	1.752	2.024	$k=5$
	(20)	(130)	(129)	(137)	(538)	
	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	

Notes: The outcome is mean number of children ever born by father's occupation (origin) and husband's 1962 occupation (destination) among wives aged 42 to 61 years in March 1962 who were currently living with their husband in the OCG sample. Numbers in parentheses are number of respondents for each cell. Total sample is $R = 5,958$.

Table 3: Comparing Nonlinear Effects Using Data from Sobel (1981)

	Origin Nonlinear Effects					Destination Nonlinear Effects				
	$\widetilde{\alpha}_1$	$\widetilde{\alpha}_2$	$\widetilde{\alpha}_3$	$\widetilde{\alpha}_4$	$\widetilde{\alpha}_5$	$\widetilde{\beta}_1$	$\widetilde{\beta}_2$	$\widetilde{\beta}_3$	$\widetilde{\beta}_4$	$\widetilde{\beta}_5$
DRM	0.085	−0.078	0.014	−0.133	0.112	0.150	−0.139	0.025	−0.235	0.199
Ratio	0.361	0.361	0.361	0.361	0.361	0.639	0.639	0.639	0.639	0.639
L-ODM model	0.122	−0.119	−0.045	−0.038	0.081	0.105	−0.097	0.069	−0.265	0.189
Ratio	0.538	0.552	−1.918	0.125	0.300	0.462	0.448	2.918	0.875	0.700

Notes: Nonlinear effects for origin and destination are shown for the DRM and the L-ODM model using Sobel's (1981) fertility data. Nonlinear effects for origin and destination, respectively, are expressed as deviations from their respective overall levels and linear effects. Ratios for the origin nonlinear effects are calculated as $\widetilde{\alpha}_i/(\widetilde{\alpha}_i + \widetilde{\beta}_{[j=i]})$, while those for the destination nonlinear effects are calculated as $\widetilde{\beta}_j/(\widetilde{\alpha}_{[i=j]} + \widetilde{\beta}_j)$. For example, the ith origin nonlinear effect is divided by the sum of the ith and jth origin and destination nonlinear effects. Total sample is $R = 5,958$.

Table 4: Biased Mobility Linear Effects Using the DRM

True DGP:	γ	-1.000	-0.500	-0.250	-0.010	0.250	0.500	1.000
Origin Effects:	$\widehat{\alpha}$	-0.115	-0.115	-0.115	-0.115	-0.115	-0.115	-0.115
	$\widehat{\alpha}^2$	0.041	0.041	0.041	0.041	0.041	0.041	0.041
	$\widehat{\alpha}^3$	0.013	0.013	0.013	0.013	0.013	0.013	0.013
	$\widehat{\alpha}^4$	0.016	0.016	0.016	0.016	0.016	0.016	0.016
Destination Effects:	$\widehat{\beta}$	-0.202	-0.202	-0.202	-0.202	-0.202	-0.202	-0.202
	$\widehat{\beta}^2$	0.073	0.073	0.073	0.073	0.073	0.073	0.073
	$\widehat{\beta}^3$	0.024	0.024	0.024	0.024	0.024	0.024	0.024
	$\widehat{\beta}^4$	0.029	0.029	0.029	0.029	0.029	0.029	0.029
Mobility Effects:	$\widehat{\gamma}$	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010	-0.010
	$\widehat{\gamma}^2$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	$\widehat{\gamma}^3$	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007
	$\widehat{\gamma}^4$	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	$\widehat{\gamma}^5$	0.003	0.003	0.003	0.003	0.003	0.003	0.003
	$\widehat{\gamma}^6$	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	$\widehat{\gamma}^7$	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	$\widehat{\gamma}^8$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001

Notes: DGP = data generating parameter. Number of origin and destination groups is set at $I = 5$ and $J = 5$, respectively, for all simulations. Sample size for each simulation is $R = 5,958$. The top row indicates the true data generating parameter for the mobility linear effect for a given simulated data set, while the remaining rows give the corresponding DRM estimates. Shaded column indicates when the DRM recovers the true mobility linear effect. The true linear effect of destination is given by $\beta = (\gamma+\beta)-\gamma$, or $-0.213-\gamma$. The true linear effect of origin is given by $\alpha = (\alpha+\beta)-(\gamma+\beta)+\gamma$, or $-0.104+\gamma$. The remaining data generating parameters are given by $\mu = 2.3171$, $\alpha^2 = 0.0411$, $\alpha^3 = 0.0136$, $\alpha^4 = 0.0161$, $\beta^2 = 0.0730$, $\beta^3 = 0.0242$, $\beta^4 = 0.0285$, $\gamma^2 = -0.0012$, $\gamma^3 = -0.0074$, $\gamma^4 = 0.0004$, $\gamma^5 = 0.0027$, $\gamma^6 = -0.0022$, $\gamma^7 = 0.0017$, $\gamma^8 = -0.0008$. All data generating parameters are derived from fertility data used by Sobel (1981). For simplicity, and without loss of generality, we assume no random error.

Table 5: Sensitivity of Estimated Mobility Linear Effect
to Values of True Origin and Destination Linear Effects

	True DGP			DRM Estimates			Bias of $\widehat{\gamma}$				
	α	β	γ	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$	$(\beta)(\widehat{w}_o)$	$-$	$(\alpha)(\widehat{w}_d)$	$=$	$\widehat{\gamma}-\gamma$
Varying α:	0.500	0.250	0.250	0.375	0.375	0.125	(0.250)(0.500)	$-$	(0.500)(0.500)	$=$	0.125
	0.250	0.250	0.250	0.250	0.250	0.250	(0.250)(0.500)	$-$	(0.250)(0.500)	$=$	0.000
	0.050	0.250	0.250	0.150	0.150	0.350	(0.250)(0.500)	$-$	(0.050)(0.500)	$=$	0.100
	0.000	0.250	0.250	0.125	0.125	0.375	(0.250)(0.500)	$-$	(0.000)(0.500)	$=$	0.125
	-0.050	0.250	0.250	0.100	0.100	0.400	(0.250)(0.500)	$+$	(0.050)(0.500)	$=$	0.150
	-0.250	0.250	0.250	0.000	0.000	0.500	(0.250)(0.500)	$+$	(0.250)(0.500)	$=$	0.250
	-0.500	0.250	0.250	-0.125	-0.125	0.625	(0.250)(0.500)	$+$	(0.500)(0.500)	$=$	0.375
Varying β:	0.250	0.500	0.250	0.375	0.375	0.375	(0.500)(0.500)	$-$	(0.250)(0.500)	$=$	0.125
	0.250	0.250	0.250	0.250	0.250	0.250	(0.250)(0.500)	$-$	(0.250)(0.500)	$=$	0.000
	0.250	0.050	0.250	0.150	0.150	0.150	(0.050)(0.500)	$-$	(0.250)(0.500)	$=$	-0.100
	0.250	0.000	0.250	0.125	0.125	0.125	(0.000)(0.500)	$-$	(0.250)(0.500)	$=$	-0.125
	0.250	-0.050	0.250	0.100	0.100	0.100	$(-0.050)(0.500)$	$-$	(0.250)(0.500)	$=$	-0.150
	0.250	-0.250	0.250	0.000	0.000	0.000	$(-0.250)(0.500)$	$-$	(0.250)(0.500)	$=$	-0.250
	0.250	-0.500	0.250	-0.125	-0.125	-0.125	$(-0.500)(0.500)$	$-$	(0.250)(0.500)	$=$	-0.375

Notes: Number of origin and destination groups is set at $I = 5$ and $J = 5$, respectively, for all simulations. Sample size for each simulation is $R = 5,958$. Shaded rows indicate that the DRM recovers the true mobility linear effect. All nonlinear effects are zero in the data generating model. For simplicity, and without loss of generality, we assume no random error. Note that, because there are no nonlinear effects included in the data generating parameters, that the origin and destination weights generated by the DRM are both 0.500. The bias arises because the underlying origin and destination linear effects do not obey the proportionality constraints of these weights. For example, in the top row, the true origin linear effect is $\alpha = 0.500$ and the true destination linear effect is 0.250, so the true weights needed to recover these effects are $w_o = 0.500/(0.250 + 0.500) = 0.667$ and $w_d = 0.250/(0.250 + 0.500) = 0.333$. Yet the estimated origin weights are simply $\widehat{w}_o = \widehat{w}_d = 0.500$.

Table 6: Typology of Models for a Mobility Table

Origin	Destination	Mobility	
$f(O)$	$f(O, D)$	$f(O, M)$	Origin
	$f(D)$	$f(D, M)$	Destination
		$f(M)$	Mobility

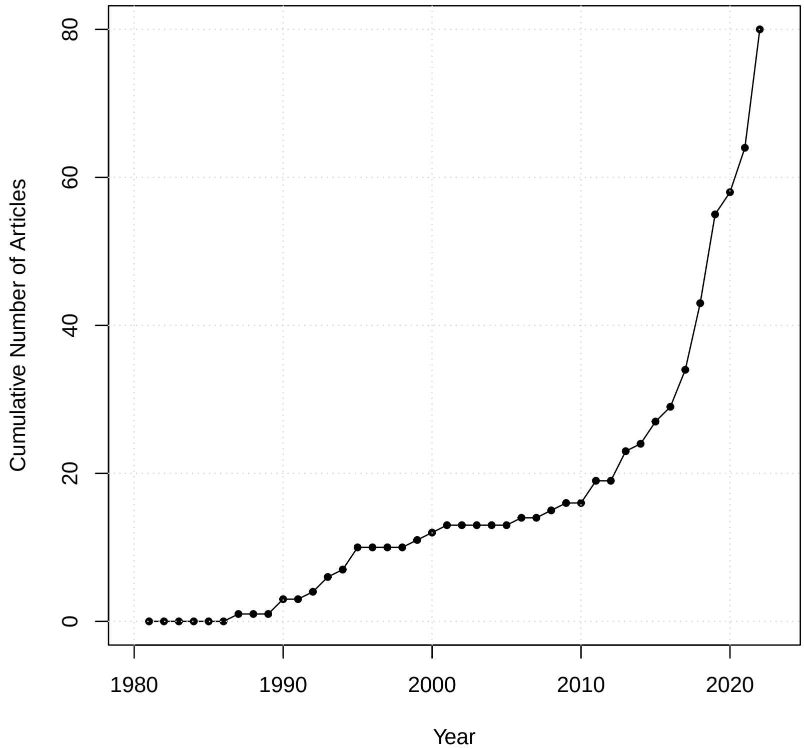
Notes: This table presents a typology of possible models given a mobility table or, more generally, data with class origin, class destination, and class mobility variables. In the table above, O denotes class origin, D class destination, and M class mobility, with $f(\cdot)$ defining some general function. Preferred models are highlighted in gray. Note that these functions are all distinct from a conventional mobility effects model, which has the form of $f(O^*, D^*, M^*)$, where O^*, D^*, and M^* are unobserved causal factors proxied by O, D, and M.

Table 7: Recommended Procedure for Analyzing Social Mobility Data

Step	**Description**	**Reference(s)**
1	Fit the Structural and Dynamic Inequality (SDI) model	Section III.2 (see Eq. 15)
2	Interpret the Social Structure (ST) and Social Mobility (SM) slopes, the ST and SM curves, and the Social Structure and Social Mobility matrices	Section III.4, Section III.5
3	Interpret the overall and adjusted comparative mobility curves	Section III.6 (see Eqs. 16, 17)
4	If one desires to partially identify unique "effects," invoke explicit assumptions about the size, sign, or shape to bound the origin, destination, or mobility "effects"	Section II.1, Section II.2
5	Provide a range of bounding strategies and emphasize the inherently tentative, non-falsifiable nature of the conclusions, as well as the strong assumptions required to interpret these quantities in terms of formal counterfactuals	Section II.2, Online Appendix F

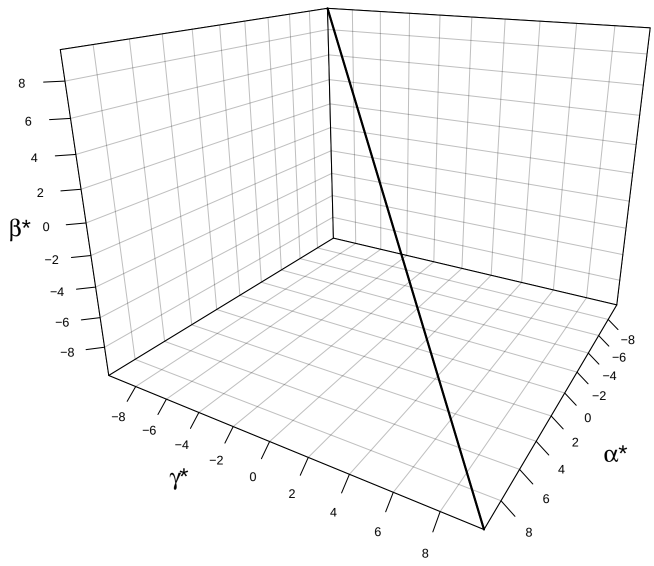
Figures

Figure 1: The Rise of the DRM



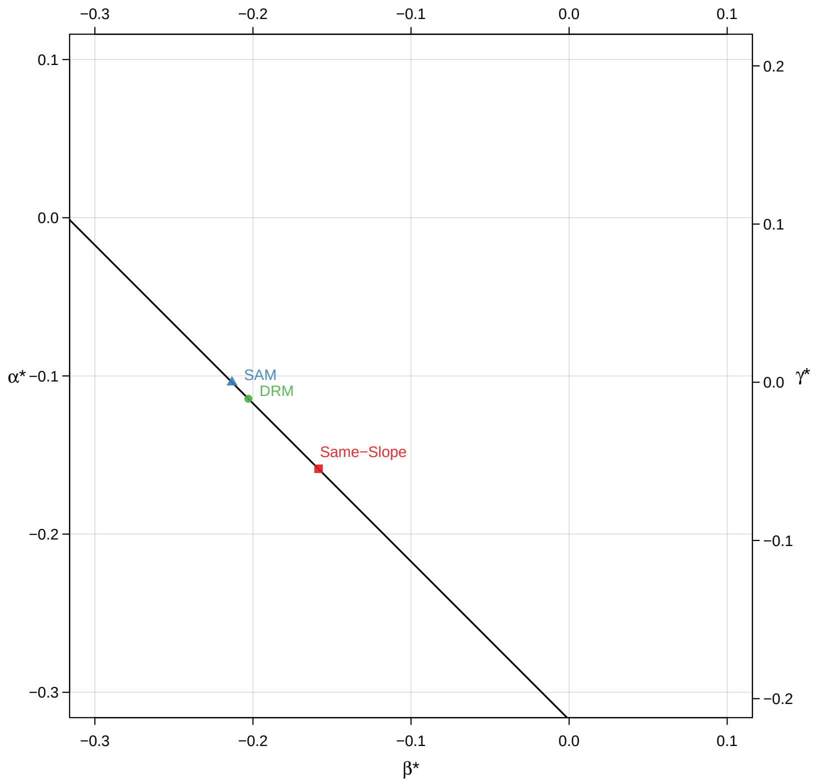
Notes: This line plot shows, by year, the cumulative number of articles using the DRM since Sobel's (1981) publication. Online Appendix D contains the full list of publications that make up the data for this graph.

Figure 2: 3D Canonical Solution Line



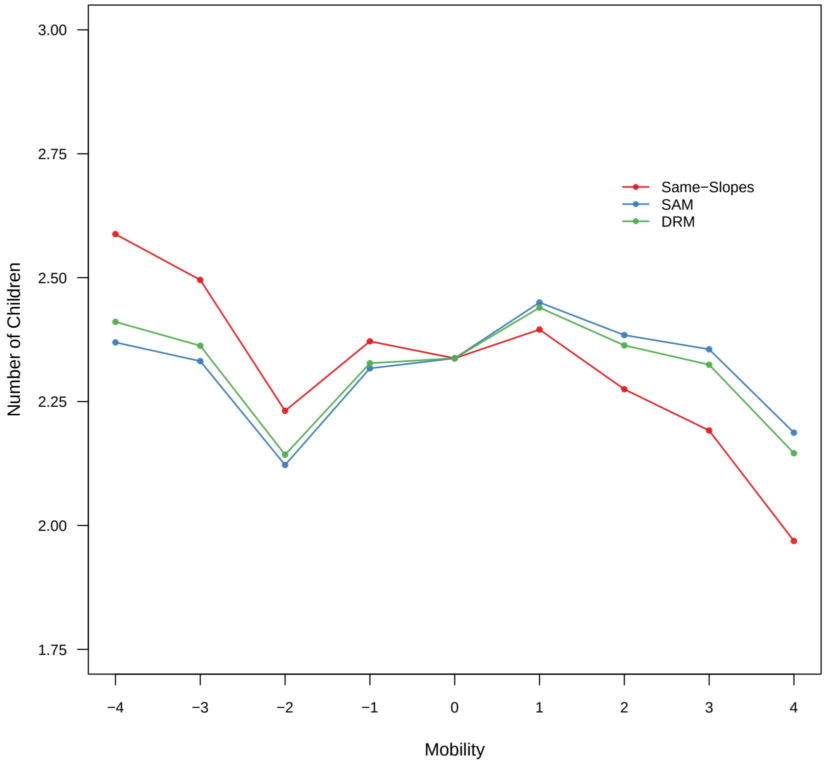
Notes: This graph shows the canonical solution line based on Sobel's (1981) fertility data. All possible combinations of origin, destination, and mobility slopes, denoted by α^*, β^*, and γ^*, respectively, lie on this line. Algebraically, the line is equivalent to the set of equations $\alpha^* = \alpha + \nu$, $\beta^* = \beta - \nu$, $\gamma^* = \gamma + \nu$, where α, β, and γ are the true origin, destination, and mobility slopes and ν is some scalar. Any given constrained model is equivalent to selecting a particular value of ν.

Figure 3: Point Identification of Mobility Effects Models: 2D Canonical Solution Line



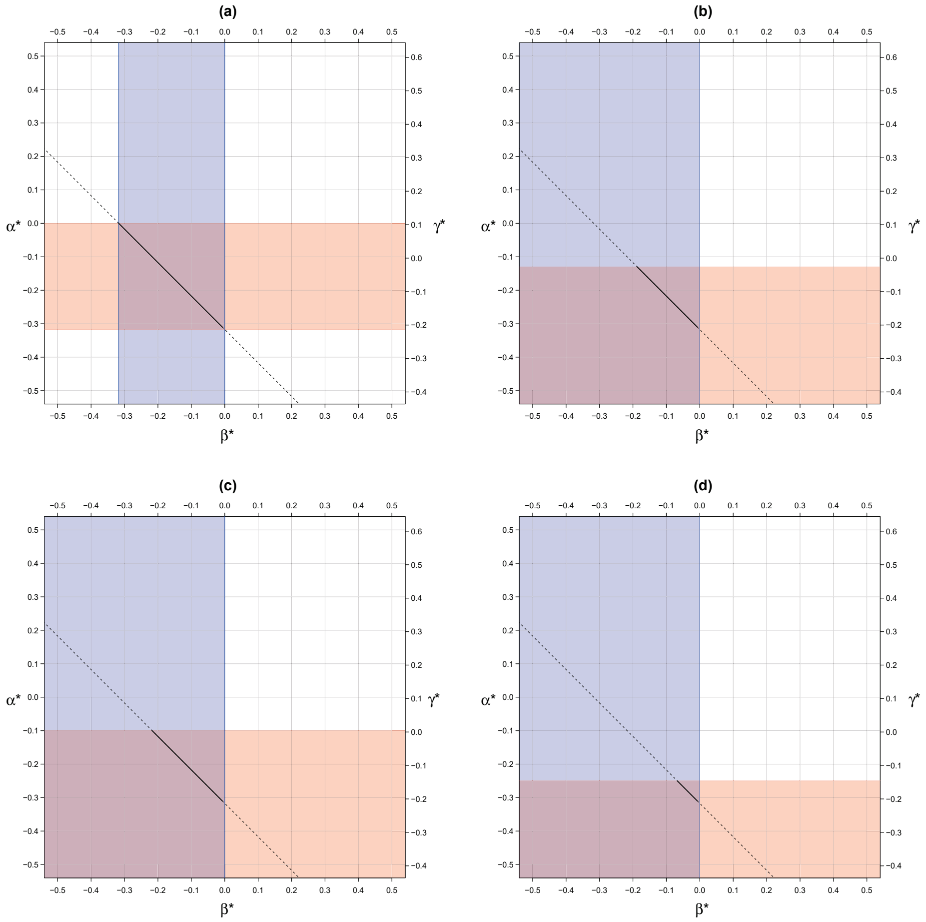
Notes: This graph shows the 2D canonical solution line based on Sobel's (1981) fertility data. Solution line is identical to that shown in Figure 2, but flattened to two dimensions. Points indicate estimated linear effects for the Square Additive Model (SAM), Diagonal Reference Model (DRM), and under the assumption that origin and destination have the same slope (Same-Slopes).

Figure 4: Point Identification of Mobility Effects Models



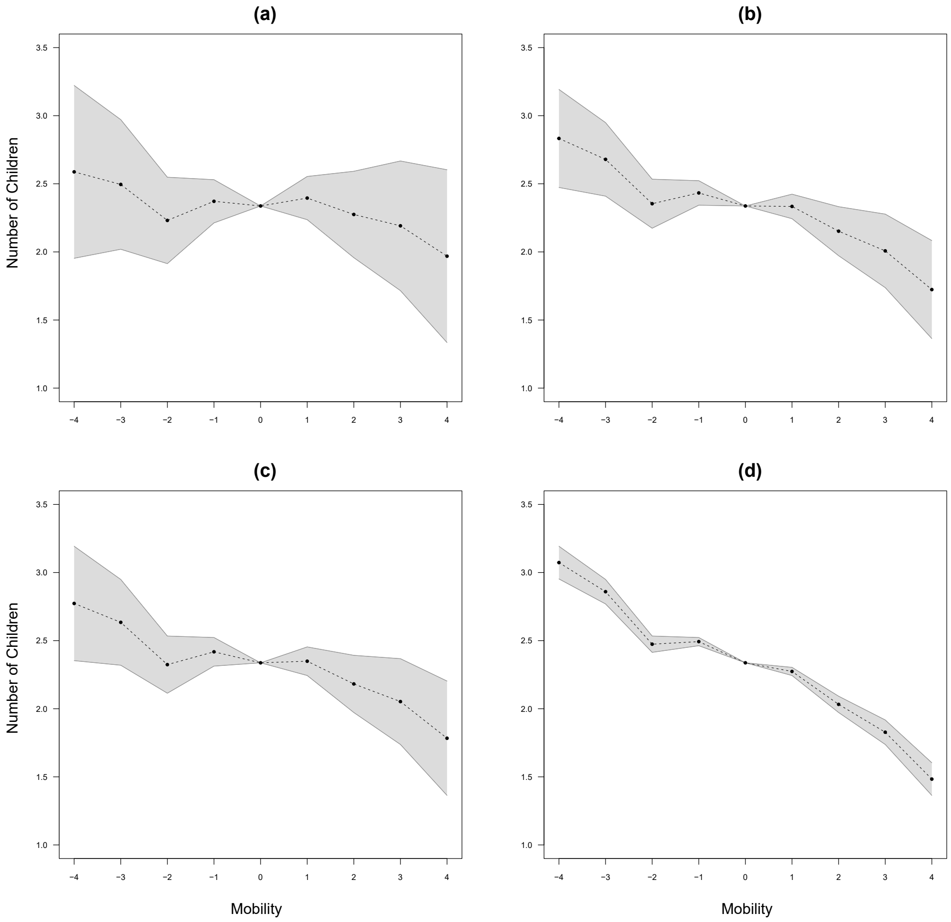
Notes: This graph shows the mobility effects for the Square Additive Model (SAM), Diagonal Reference Model (DRM), and under the same-slopes for origin and destination constraint (Same-Slopes). Data are based on Sobel (1981).

Figure 5: Partial Identification of Mobility Effects Models:
2D Canonical Solution Line



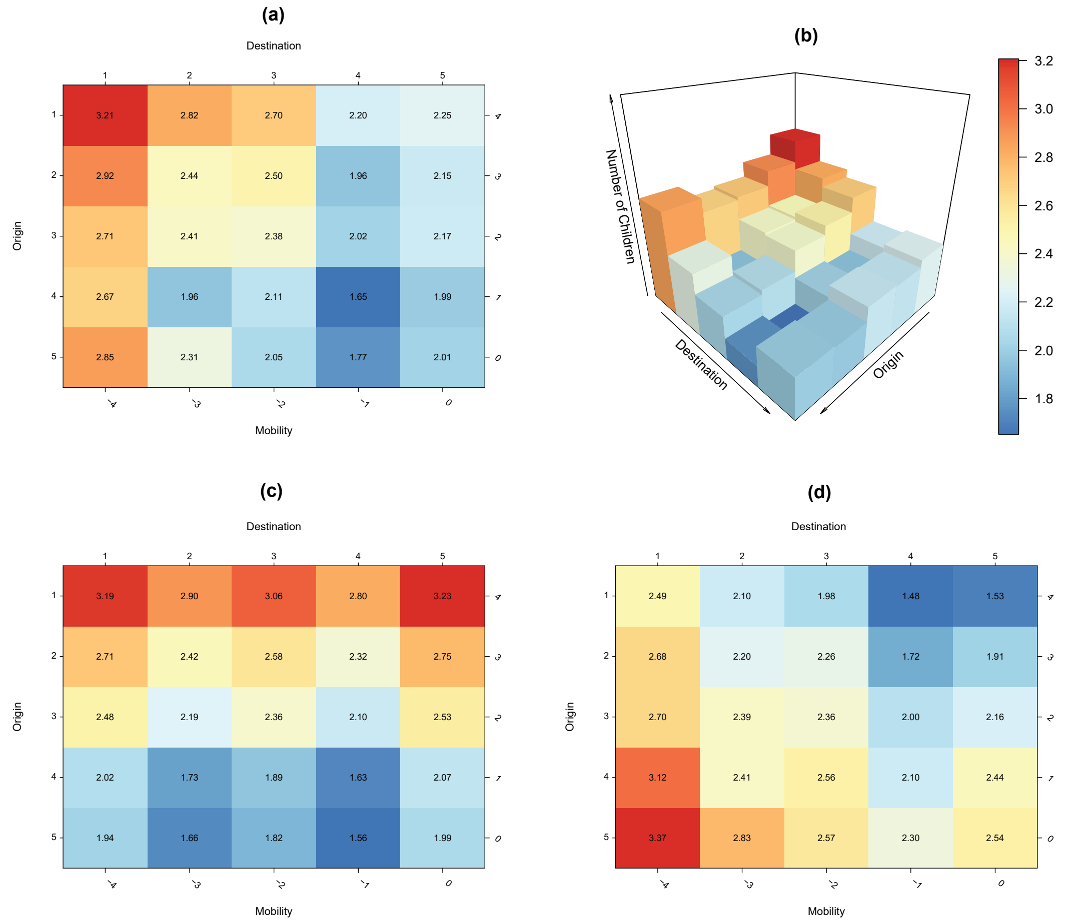
Notes: Panel (a) shows bounds on the canonical solution line using the same-sign assumption for class origin and destination. Panel (b) shows bounds on the canonical solution line under the assumption of a monotonically downward effect for class origin with respect to the quadratic component of the origin nonlinear effects as well as a same-sign assumption for class destination. Panel (c) shows bounds on the canonical solution line under the assumption of a monotonically decreasing origin effect for lower-class groups (1, 2, and 3) in addition to a same-sign assumption for class destination. Lastly, panel (d) shows bounds on the canonical solution line under the assumption of a monotonically decreasing origin effect for upper-class origin groups (3,4, and 5) as well as a same-sign assumption for class destination. Blue rectangle indicates bounds on class origin while red rectangle denotes bounds on class destination. Bold line indicates that part of the canonical solution line consistent with the data as well as the particular assumptions invoked. Data are based on Sobel (1981).

Figure 6: Partial Identification of Mobility Effects Models



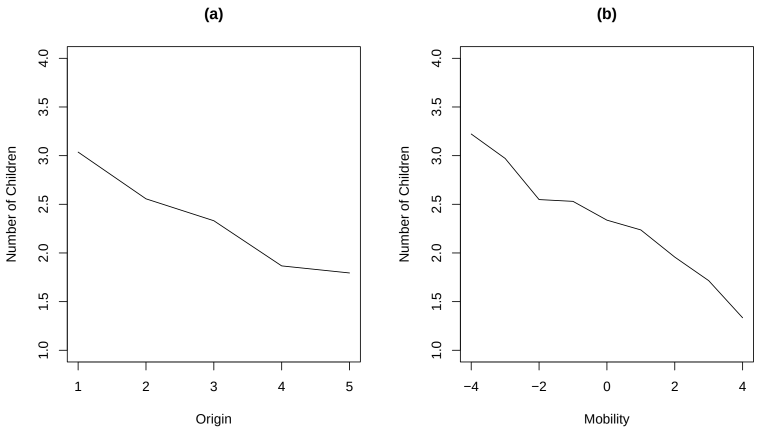
Notes: Panel (a) shows the mobility effects using the same-sign assumption for class origin and destination. Panel (b) shows bounds on the mobility effects under the assumption of a monotonically downward effect for class origin with respect to the quadratic component of the origin nonlinear effects as well as a same-sign assumption for class destination. Panel (c) shows bounds on the mobility effects under the assumption of a monotonically decreasing origin effect for lower-class groups (1, 2, and 3) in addition to a same-sign assumption for class destination. Lastly, panel (d) shows bounds on the mobility effects under the assumption of a monotonically decreasing origin effect for upper-class origin groups (3,4, and 5) as well as a same-sign assumption for class destination. Dotted line indicates the midpoint of the bounds. Data are based on Sobel (1981).

Figure 7: Social Mobility Matrices



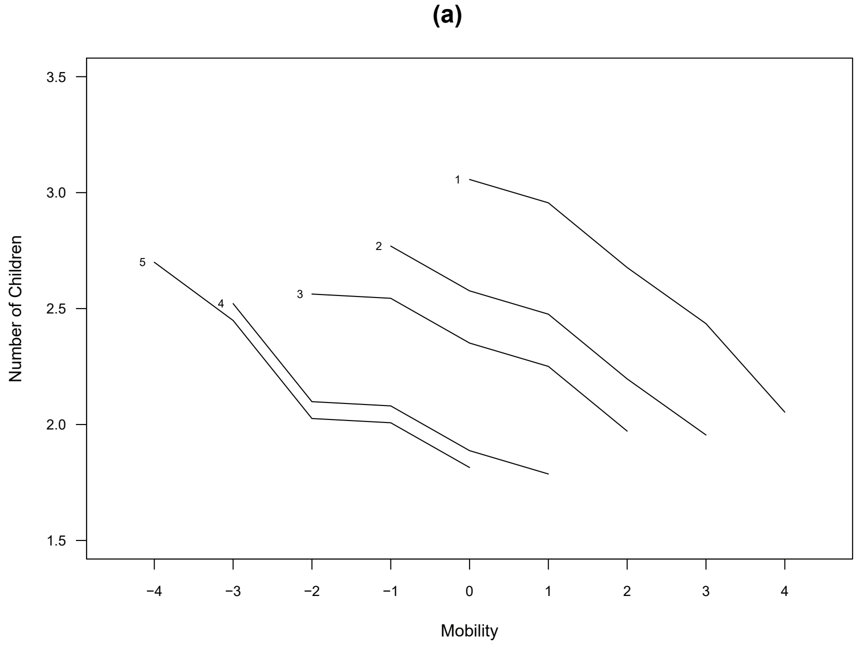
Notes: Panels (a) and (b) present 2D and 3D visualizations, respectively, of the predicted mean fertility based on the SDI model's main parameters. That is, each cell is given by $\widehat{\mu}_{ijk} = \mu + \Gamma_1(i - i^*) + \Gamma_2(k - k^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{[i+k-I]} + \widetilde{\gamma}_k$. Panel (c) shows the mean fertility based on the joint origin-destination mobility parameters, such that $\widehat{\mu}_{ijk} = \mu + \Gamma_1(i - i^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{[i+k-I]}$. Finally, panel (d) shows the mean fertility based on the joint mobility-destination mobility parameters, such that $\widehat{\mu}_{ijk} = \mu + \Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{[i+k-I]}$. Data are based on Sobel (1981).

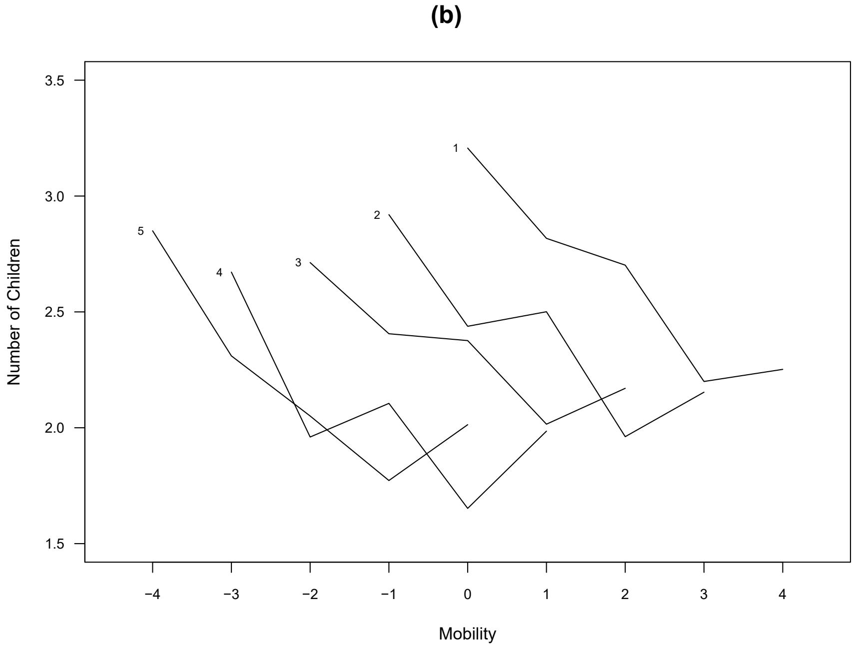
Figure 8: Social Structure and Social Mobility Curves



Notes: Panel (a) shows the social structure (ST) curve, given by $\Gamma_1(i-i^*)+\widetilde{\alpha}_i$ for class origin groups $i = 1, \ldots, I$. Panel (b) shows the social mobility (SM) curve, given by $\Gamma_2(k-k^*)+\widetilde{\gamma}_k$ for mobility groups $k = 1, \ldots, K$. $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$. Data are based on Sobel (1981).

Figure 9: Comparative Mobility Curves





Notes: Panel (a) shows the overall comparative mobility curve. Each curve is a function of $\phi_i + \Gamma_2(k - k^*) + \widetilde{\gamma}_k$ for all k in each origin group i. Panel (b) shows the adjusted comparative mobility curve. Each curve is a function of $\phi_i + \Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{i+k-I}$ for all k in each origin group i. For both curves ϕ_i is equal to $\Gamma_1(i - i^*) + \widetilde{\alpha}_i$, which is a single value for a given origin group i. Data are based on Sobel (1981).

Online Appendix A: Supplemental Tables and Figures

Table A.1: Bounding Formulas for Slopes

Origin Bounds:	$\alpha_{\min} \leq \alpha \leq \alpha_{\max}$
	$\Gamma_1 - \alpha_{\max} \leq \beta \leq \Gamma_1 - \alpha_{\min}$
	$(\Gamma_2 - \Gamma_1) + \alpha_{\min} \leq \gamma \leq (\Gamma_2 - \Gamma_1) + \alpha_{\max}$
Destination Bounds:	$\Gamma_1 - \beta_{\max} \leq \alpha \leq \Gamma_1 - \beta_{\min}$
	$\beta_{\min} \leq \beta \leq \beta_{\max}$
	$\Gamma_2 - \beta_{\max} \leq \gamma < \Gamma_2 - \beta_{\min}$
Mobility Bounds:	$(\Gamma_1 - \Gamma_2) + \gamma_{\min} \leq \alpha \leq (\Gamma_1 - \Gamma_2) + \gamma_{\max}$
	$\Gamma_2 - \gamma_{\min} \leq \beta \leq \Gamma_2 - \gamma_{\max}$
	$\gamma_{\min} \leq \gamma \leq \gamma_{\max}$

Notes: Origin, destination, and mobility slopes are α, β, and γ, respectively, with $(.)_{\min}$ and $(.)_{\max}$ denoting minimum and maximum values of the bounds. We denote $\Gamma_1 = \alpha + \beta$, $\Gamma_2 = \beta + \gamma$, $\Gamma_1 - \Gamma_2 = \alpha - \gamma$, and $\Gamma_2 - \Gamma_1 = \gamma - \alpha$.

Table A.2: Bounds Given by Setting the Sign of One or Two Slopes

Sign of One Slope	Origin	Destination	Mobility
If $\alpha \geq 0$ then:	$0 \leq \alpha < +\infty$	$-\infty < \beta \leq \Gamma_1$	$(\Gamma_2 - \Gamma_1) \leq \gamma < +\infty$
If $\alpha \leq 0$ then:	$-\infty < \alpha \leq 0$	$\Gamma_1 \leq \beta < +\infty$	$-\infty < \gamma \leq (\Gamma_2 - \Gamma_1)$
If $\beta \geq 0$ then:	$-\infty < \alpha \leq \Gamma_1$	$0 \leq \beta < +\infty$	$-\infty < \gamma \leq \Gamma_2$
If $\beta \leq 0$ then:	$\Gamma_1 \leq \alpha < +\infty$	$-\infty < \beta \leq 0$	$\Gamma_2 \leq \gamma < +\infty$
If $\gamma \geq 0$ then:	$(\Gamma_1 - \Gamma_2) \leq \alpha < +\infty$	$-\infty < \beta \leq \Gamma_2$	$0 \leq \gamma < +\infty$
If $\gamma \leq 0$ then:	$-\infty < \alpha \leq (\Gamma_1 - \Gamma_2)$	$\Gamma_2 \leq \beta < +\infty$	$-\infty < \gamma \leq 0$
Sign of Two Slopes	**Origin**	**Destination**	**Mobility**
If $\alpha \geq 0$ and $\beta \geq 0$ then:	$0 \leq \alpha \leq \Gamma_1$	$0 \leq \beta \leq \Gamma_1$	$(\Gamma_2 - \Gamma_1) \leq \gamma \leq \Gamma_2$
If $\alpha \leq 0$ and $\beta \leq 0$ then:	$\Gamma_1 \leq \alpha \leq 0$	$\Gamma_2 \leq \beta \leq 0$	$\Gamma_2 \leq \gamma \leq (\Gamma_2 - \Gamma_1)$
If $\beta \geq 0$ and $\gamma \geq 0$ then:	$(\Gamma_1 - \Gamma_2) \leq \alpha \leq \Gamma_1$	$0 \leq \beta \leq \Gamma_2$	$0 \leq \gamma \leq \Gamma_2$
If $\beta \leq 0$ and $\gamma \leq 0$ then:	$\Gamma_1 \leq \alpha \leq (\Gamma_1 - \Gamma_2)$	$\Gamma_2 \leq \beta \leq 0$	$\Gamma_2 \leq \gamma \leq 0$
If $\alpha \geq 0$ and $\gamma \leq 0$ then:	$0 \leq \alpha \leq (\Gamma_1 - \Gamma_2)$	$\Gamma_2 \leq \beta \leq \Gamma_1$	$(\Gamma_2 - \Gamma_1) \leq \gamma \leq 0$
If $\alpha \leq 0$ and $\gamma \geq 0$ then:	$(\Gamma_1 - \Gamma_2) \leq \alpha \leq 0$	$\Gamma_1 \leq \beta \leq \Gamma_2$	$0 \leq \gamma \leq (\Gamma_2 - \Gamma_1)$

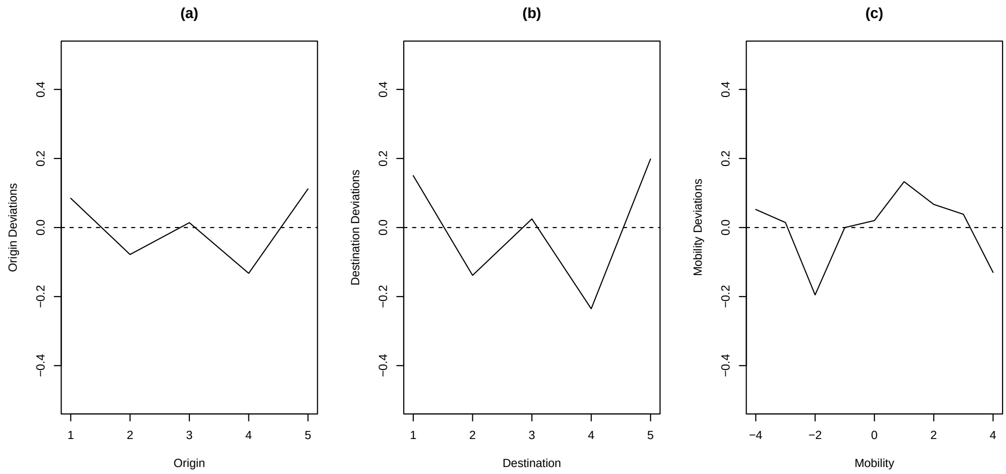
Notes: Origin, destination, and mobility slopes are α, β, and γ, respectively, with $(.)_{\min}$ and $(.)_{\max}$ denoting minimum and maximum values of the bounds. We denote $\Gamma_1 = \alpha + \beta$, $\Gamma_2 = \beta + \gamma$, $\Gamma_1 - \Gamma_2 = \alpha - \gamma$, and $\Gamma_2 - \Gamma_1 = \gamma - \alpha$.

Table A.3: Relationships among ST, SM, & Intra-Destination Slopes

Social Structure (ST) Slope:	If $\Gamma_1 = 0$,	then:	$\Gamma_2 - \Gamma_1$	$=$	Γ_2
	If $\Gamma_1 > 0$,	then:	$\Gamma_2 - \Gamma_1$	$<$	Γ_2
	If $\Gamma_1 < 0$,	then:	$\Gamma_2 - \Gamma_1$	$>$	Γ_2
Social Mobility (SM) Slope:	If $\Gamma_2 = 0$,	then:	$\Gamma_1 - \Gamma_2$	$=$	Γ_1
	If $\Gamma_2 > 0$,	then:	$\Gamma_1 - \Gamma_2$	$<$	Γ_1
	If $\Gamma_2 < 0$,	then:	$\Gamma_1 - \Gamma_2$	$>$	Γ_1
Intra-Destination Origin Slope:	If $\Gamma_1 - \Gamma_2 = 0$,	then:	Γ_1	$=$	Γ_2
	If $\Gamma_1 - \Gamma_2 > 0$,	then:	Γ_1	$>$	Γ_2
	If $\Gamma_1 - \Gamma_2 < 0$,	then:	Γ_1	$<$	Γ_2
Intra-Destination Mobility Slope:	If $\Gamma_2 - \Gamma_1 = 0$,	then:	Γ_2	$=$	Γ_1
	If $\Gamma_2 - \Gamma_1 > 0$,	then:	Γ_2	$>$	Γ_1
	If $\Gamma_2 - \Gamma_1 < 0$,	then:	Γ_2	$<$	Γ_1

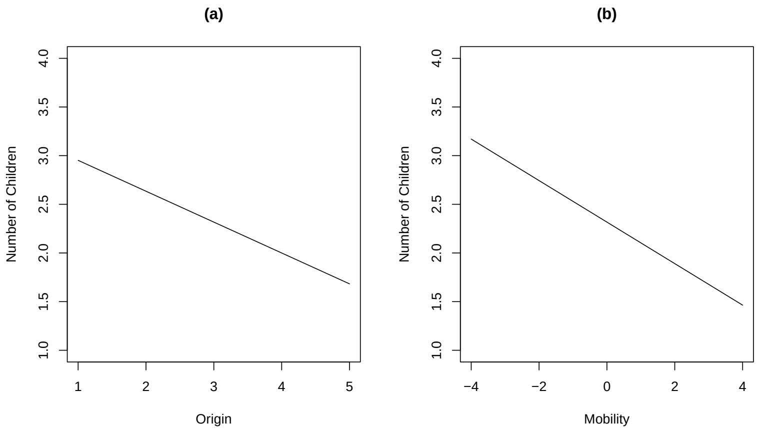
Notes: $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$.

Figure A.1: Nonlinearities for Origin, Destination, and Mobility



Notes: Panels (a), (b), and (c) show the origin, destination, and mobility nonlinearities, which are constrained to sum to zero. Data are based on Sobel (1981).

Figure A.2: Social Structure and Social Mobility Slopes



Notes: Panel (a) shows the social structure (ST) slope, while panel (b) shows the social mobility (SM) slope. $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$. Data are based on Sobel (1981).

Online Appendix B: Mobility Models & The Identification Challenge: A General Introduction

In this section, we first present a general mobility effects model that incorporates primary parameters capturing the independent effects of origin, destination, and mobility, along with additional cell-specific parameters that represent heterogeneity or interactions among these dimensions. We then introduce a reparameterized version of the conventional model that explicitly distinguishes linear from nonlinear effects. This reparameterization facilitates a deeper understanding of the inherent limitations in identifying mobility effects while also laying the groundwork for the various approaches to analyzing mobility effects we discuss in subsequent sections.

In the conventional mobility effects literature, researchers pursue the identification of unique origin, destination, and mobility effects despite the mathematical dependency among these dimensions. While not explicitly articulated, this pursuit implies a conceptual distinction between observed positions on a mobility table, namely, origin, destination, and mobility, and the underlying causal mechanisms they represent. That is, although destination (D) is mathematically defined as the sum of origin (O) and mobility (M), the bundles of causal mechanisms they proxy for, which we denote as O^*, D^*, and M^*, are theoretically distinct and capable of varying independently. Origin effects (O^*) might represent parental economic resources, cultural socialization, or educational guidance; destination effects (D^*) could capture workplace authority relations, professional network benefits, or class-based consumption opportunities; and mobility effects (M^*) might reflect distinct processes such as status anxiety, reference group changes, or psychological adaptation to class transitions. This distinction supplies the rationale for attempting to identify unique effects of origin, destination, and mobility, notwithstanding their deterministic mathematical relationship. More formally, let O^*, D^*, and M^* denote underlying bundles of causal mechanisms that are allowed to vary freely from each other such that $D^* \neq O^* + M^*$. Mobility effects analysis, as commonly used in the literature, can generally be understood as any analysis using functions of the form $Y = f(O^*, D^*, M^*) + \epsilon$, where ϵ is a normally distributed error term with a mean of zero. However, because O^*, D^*, and M^* are typically unobserved, the observed dimensions of the mobility table, O, D, and M, which have the natural relationship $D = O + M$, are substituted for the underlying causal variables O^*, D^*, and M^* (cf. Bijlsma et al. 2017: 722-724; Clogg 1982; Heckman and Robb 1985). As we show in later sections, it is only under very strong assumptions that one can extract unique "effects" using a conventional analysis of mobility effects.

More specifically, suppose we have data collected on individuals indexed from $r = 1, \ldots, R$, where R is the total number of respondents. Additionally, suppose we have data collected on the underlying causal factors O^*, D^*, M^*, which are coded as categorical variables with levels indexed by $l = 1, \ldots, L$, $p = 1, \ldots, P$, and $n = 1, \ldots, N$, respectively.[46] The mobility effects model can thus be specified as follows:

$$Y = f(O^*, D^*, M^*) + \epsilon = \mu^* + \alpha_l^* + \beta_p^* + \gamma_n^* + \eta_{lpn}^* + \xi_{rlpn}^*, \tag{B.1}$$

where μ^* is the intercept (or overall mean); α_l^*, β_p^*, γ_n^* denote the lth, pth, nth levels of the underlying causal factors for origin, destination, and mobility, respectively; η_{lpn}^* is an additional (orthogonal) term denoting interactions among the underlying factors; and ξ_{rlpn}^* is an individual-level, normally distributed error term with a mean of zero. If one somehow had access to these underlying factors, then, under standard assumptions of consistency and no interference between units, positivity and overlap, and conditional ignorability, we could use Equation B.1 to estimate $\mathbb{E}[Y^{o^*d^*m^*}]$,

[46] For simplicity, and without loss of generality, we will assume that the origin and destination categories are of equal width.

the expected value of the counterfactual outcome if we were to set O^* to some value o^*, D^* to some value d^*, and M^* to some value m^*, each of which, as noted above, are allowed to freely vary. The estimand here is the difference in expected outcomes for two hypothetical individuals who share the same origin o^* and the same destination d^* but differ in their mobility factor m^*. Formally, one can write this "mobility effect" as:

$$\mathbb{E}\big[Y^{(o^*,d^*,m^*)}\big] \;-\; \mathbb{E}\big[Y^{(o^*,d^*,m'^*)}\big],$$

where o^* and d^* are held fixed at the same levels in both expectations, and $m^* \neq m'^*$ represents two different possible mobility "treatments." Under the usual assumptions (consistency, no interference, positivity, ignorability, as well as the assumption that the causal directions are appropriately specified), this difference captures how much the outcome Y would change if origin and destination were fixed to the same values but mobility status were fixed to different values. This is the estimand that the rapidly growing literature on social mobility effects, which overwhelmingly uses the DRM (see Online Appendix C), seeks to identify. The methodological challenges we discuss in the section "The Diagonal Reference Model" apply to this estimand but not necessarily to other causal estimands, such as those identified in the section "A Paradigm Shift".

In practice, one typically does not have access to the underlying bundles of causal mechanisms in Equation B.1. Instead of treating them as dimensions of observed positionality on a mobility table, O, D, M, as noted above, are used as proxies for O^*, D^*, and M^*.[47] Specifically, suppose we have a set of categorical variables for $i = 1, \ldots I$ origin groups, $j = 1, \ldots, J$ destination groups, and $k = j - i + I, \ldots K$ mobility groups. The mobility effects model using origin, destination, and mobility as proxies can accordingly be specified using what we call the *Classical Origin-Destination-Mobility (C-ODM) model*:

$$Y = f(O^*, D^*, M^*) + \epsilon = \mu + \alpha_i + \beta_j + \gamma_k + \eta_{ijk} + \xi_{rijk}, \tag{B.2}$$

where μ is the intercept (or overall mean); α_i, β_j, γ_k denote the ith, jth, kth observed levels of origin, destination, and mobility, respectively; η_{ijk} is an additional (orthogonal) term denoting interactions; and ξ_{rijk} is an individual-level, normally-distributed error term with a mean of zero. To reiterate, Equation B.2 is based on the implicit assumption that O, D, and M (and their respective indices) can be treated as surrogates for O^*, D^*, and M^* (and their respective indices). To simplify the exposition, we will accordingly refer to α_i, β_j, and γ_k as the "true" origin, destination, and mobility effects, but the reader should keep in mind that this is shorthand for referring to α^*_l, β^*_p, γ^*_n (for a similar point, see Fosse and Winship 2019a).

Unfortunately, the basic mobility effects model outlined in Equation B.2 suffers from a fundamental identification problem that that goes beyond the identification problem common to all linear models using categorical variables as inputs.[48] This problem was vividly illustrated by the

[47] An alternative strategy is to shift away from modeling general "effects" and instead focus on examining the effects of specific variables that are thought to capture particular origin-, destination-, or mobility-related processes. For example, rather than attempting to model an omnibus "mobility effect" on, say, voting behavior or political preferences (e.g., Clifford and Heath 1993; De Graaf et al. 1995), one might examine how specific mobility-related events, such as a spell of unemployment or changes in job tasks, affect the likelihood of voting for a particular political party (e.g., Turner and Ryan 2023; Wiertz and Rodon 2021). However, because this approach focuses on understanding the effects of particular mechanisms rather than global origin, destination, and mobility effects, it may be seen as a shift away from mobility effects analysis as it has been traditionally understood in the literature.

[48] The common identification problem is that, with an intercept in the model, there is one more level than can be estimated for the origin, destination, and mobility effect. Although common, interpretation errors can ensue: For example, in a related literature on APC models, it has been shown that for some estimators seemingly trivial changes in coding schemes, such as the level used as the reference category, can generate dramatically different results (Fosse

sociologist Hubert Blalock (1966: 53), who posed the following thought experiment: "Suppose an unscrupulous demon were to perform certain legitimate mathematical manipulations, presenting to us some new equations with different numerical values for the slopes. Could we ever discover the hoax?" Unfortunately, with respect to the analysis of mobility effects, the answer is in the negative: the linear effects are not identified in conventional mobility models, and estimates are compatible with an infinite range of possible values (Blalock 1966, 1967; Duncan 1966; Sobel 1981, 1985).[49] Intuitively, this is simply because there is not enough information to identify all three linear effects from the data alone (for a related discussion and proofs, see Fosse and Winship 2018).

It is worth emphasizing that the identification challenge in mobility research shares important similarities with the classic age-period-cohort (APC) problem, in which $\text{Age} + \text{Cohort} = \text{Period}$ creates a linear dependency that cannot be resolved with data alone. However, APC analyses typically rely on a Lexis table indexing temporally-based dimensions, namely, age, historical period, and birth cohort, which serve as proxies for life-cycle, generational, and period-based causal processes (e.g., see Fosse and Winship 2019b). By contrast, mobility research is grounded in structural dimensions, namely, class origin, class destination, and social mobility, that proxy class- and movement-based causal processes. These different substantive applications have informed the history of model development in both domains. Despite these distinct substantive interpretations, however, both APC and origin-destination-mobility models confront the same underidentification challenge: neither set of "effects" can be uniquely disentangled without additional, often strong, assumptions.

An alternative formulation of the C-ODM model (see Equation B.2) helps clarify the nature of the identification problem. By orthogonalizing the linear from the nonlinear terms, we can specify what we call the *Linearized Origin-Destination-Mobility (L-ODM) model*:

$$\mu_{rijk} = \mu + \alpha(i - i^*) + \beta(j - j^*) + \gamma(k - k^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j + \widetilde{\gamma}_k + \eta_{ijk} + \xi_{rijk} \tag{B.3}$$

where the asterisks denote midpoint or referent indices $i^* = (I + 1)/2$, $j^* = (J + 1)/2$, and $k^* = (K + 1)/2$; α, β, and γ denote the linear effects of origin, destination, and mobility, respectively; and $\widetilde{\alpha}$, $\widetilde{\beta}$, and $\widetilde{\gamma}$ represent the origin, destination, and mobility nonlinear effects, respectively; η_{ijk} is, as before, an additional (orthogonal) term denoting interactions; and ξ_{rijk} is a normally-distributed individual-level error term with a mean of zero. To identify the levels of the parameters given the inclusion of the intercept, sum-to-zero constraints are applied to the linear and nonlinear parameters.

The C-ODM and L-ODM models are equivalent representations of class data grouped by origin, destination, and mobility in the sense that a model fitted using either specification will result in the same predicted values of the outcome.[50] However, the L-ODM model has a significant advantage over the C-ODM model. Due to the linear dependence among origin, destination, and mobility, as well as the fact that origin, destination, and mobility parameters combine slopes with deviations, even after applying sum-to-zero constraints, in general no parameters are identified in the C-ODM model other than the overall mean (cf. Fosse and Winship 2018).[51] By contrast, after applying the

and Winship 2018). In the discussion that follows, we will assume that sum-to-zero constraints are applied, such that $\sum_{i=1}^{I} \alpha_i = \sum_{j=1}^{J} \beta_j = \sum_{k=1}^{K} \gamma_k = 0$, with the last category of the origin, destination, and mobility variables dropped.

[49] This does not mean, however that, based on substantive or theoretical knowledge, one would necessarily believe that all values are equally valid (e.g., see Fosse and Winship 2019b).

[50] More technically and precisely, the two models are equivalent in that they span the same linear subspace of the data.

[51] However, if I, J, and/or K are odd, then under conventional "normalization" assumptions the corresponding

sum-to-zero constraints, only the three linear effects in the L-ODM model remain unidentified.[52] This greatly simplifies the nature of the identification problem and, as we show later, allows one to use graphical tools for visualizing and partially identifying the parameters of a mobility effects model. Moreover, as we demonstrate in the next section, the L-ODM can be used to clarify the assumptions underlying the current wave of studies on mobility effects.

mean parameters $\alpha_{(I+1)/2}$, $\beta_{(J+1)/2}$, and/or $\gamma_{(K+1)/2}$ will be identified (see Smith 2021 for a similar point).

[52]We elaborate on this property later when we discuss the bounding approach to mobility effects models.

Online Appendix C: Summarizing Variability on the Mobility Table

In this appendix we outline models for describing structural and dynamic inequalities in a mobility table. We first present three different models for summarizing variation in a mobility based on re-indexing the L-ODM model by origin-destination, destination-mobility, and origin-mobility, respectively. We then show how, using the logical of omitted variable bias and matrix algebra, how these models clarify what is actually estimated when fitting all three possible one-factor models and all three possible two-factor models. These results show that, in general, that for describing patterns of mobility either a model of the form $Y = f(D)$ or $Y = f(O, M)$ is to be preferred. Throughout this appendix we let $i = 1, \ldots, I$ index the origin groups, $j = 1, \ldots, J$ the destination groups, and $k = 1, \ldots, K$ the mobility groups, where $k = j - i + I$ and $K = I + J - 1$. As well, we let $r = 1, \ldots, R$ index the respondents (i.e., individuals) in the data set.

1. Three Models for Describing Dynamic and Structural Inequality

As noted in the main text, instead of attempting to identify unique or "pure" effects, one can use the L-ODM model to identify structural and dynamic processes operating on a mobility table. The key insight is that we can project the three-dimensional (unidentified) L-ODM model onto a two-dimensional surface (i.e., a mobility table) by exploiting the fact that mobility, origin, and destination are linearly related. Because there are three different ways to index a mobility table (i.e., origin-mobility, origin-destination, destination-mobility), there are three distinctly different models for describing patterns on a mobility model.[53] We outline each of these models below. Although each model is indexed by two dimensions, because each model contains parameters for all three dimensions, we will refer to them as "three-factor" models.

The first model is based on taking the L-ODM model and re-specifying it as an origin-mobility model. Note that, given an origin-destination mobility table, $j = i + k - I$ and $J = K - I + 1$. Substituting for j and J in the L-ODM model and rearranging terms leads to what we call the *Structural and Dynamic Inequality model* or the *SDI model* for short:

$$\mu_{rijk} = f(O, M) = \mu + \Gamma_1(i - i^*) + \Gamma_2(k - k^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{[i+k-I]} + \widetilde{\gamma}_k + \eta_{i[i+k-I]k} + \xi_{ri[i+k-I]k}, \quad \text{(C.1)}$$

where $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$, or the social structure (ST) slope and the social mobility (SM) slope, respectively. As a result of the substitution of the sum of the origin and mobility indices for the destination indices (that is, $j = i + k - I$ and $J = K - I + 1$), the outcome is simply a function of origin, indexed by i, with corresponding parameters representing structural inequalities, and mobility, indexed by k, with corresponding parameters representing dynamic inequalities. This model is identified (i.e., the design matrix is of full rank) as it does not contain a separate linear term for destination, which is instead combined with the origin and mobility linear terms, respectively.

The second model is based on expressing the parameters of the L-ODM model in terms of origin and destination. Note that, given an origin-destination mobility table, $k = j - i + I$ and $K = I + J - 1$. Substituting for k and K in the L-ODM model and rearranging terms results in what we call the *Intra-Destination Differences and Structural Inequality model* or, for short, the *Diff-SI model*:

$$\mu_{rijk} = f(O, D) = \mu + (\Gamma_1 - \Gamma_2)(i - i^*) + \Gamma_2(j - j^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j + \widetilde{\gamma}_{j-i+I} + \eta_{ij[j-i+I]} + \xi_{rij[j-i+I]}, \quad \text{(C.2)}$$

where $\Gamma_1 - \Gamma_2 = (\alpha + \beta) - (\gamma + \beta) = \alpha - \gamma$ and $\theta_2 = \gamma + \beta$. The difference $\Gamma_1 - \Gamma_2$ in Equation C.2 is a slope of differences within the class destination, while Γ_2 is simply the total realized mobility slope from the SDI model, but indexed by destination ($j = 1, \ldots, j = J$) instead of by mobility

[53] Note that these models take the general form of $Y = f(O, D)$, $Y = f(D, M)$, and $Y = f(O, M)$.

levels ($k = 1, \ldots, k = K$). Similar to the SDI model, the Diff-SI model is identified because it does not contain a unique linear term for social mobility, which is instead absorbed into the origin and destination linear terms.

Finally, the third logically possible model entails expressing the parameters of the L-ODM model in terms of destination and mobility. Note that, given an origin-destination mobility table, $i = j - k + I$ and $I = K - J + 1$. Substituting for i and I in the L-ODM model and rearranging terms results in what we call the *Dynamic Inequality and Intra-Destination Differences model* or, for short, the *DI-Diff model*:

$$\mu_{rijk} = f(D, M) = \mu + \Gamma_1(j - j^*) + (\Gamma_2 - \Gamma_1)(k - k^*) + \widetilde{\alpha}_{j-k+(K-J+1)} + \widetilde{\beta}_j + \widetilde{\gamma}_k + \eta_{[j-k+(K-J+1)]jk} + \xi_{r[j-k+(K-J+1)]jk}, \quad \text{(C.3)}$$

where $\Gamma_2 - \Gamma_1 = (\gamma + \beta) - (\alpha + \beta) = \gamma - \alpha$ and $\Gamma_1 = \alpha + \beta$. The difference $\Gamma_2 - \Gamma_1$ in Equation C.3 is an overall slope of differences within destination classes, while Γ_1 is simply the SI slope from the SDI model, but indexed by class destination ($j = 1, \ldots, j = J$) instead of class origin ($i = 1, \ldots, i = I$). Similar to the previous two models, the DI-Diff model is identified because it does not include a separate linear term for origin, which is instead absorbed into the destination and mobility linear terms.

All three models outlined above provide the same estimates of the intercept and the origin, destination, and mobility nonlinearities. However, unlike the SDI model, the slopes indexed by origin and mobility in Equations C.2 and C.3, respectively, are what can be deemed "synthetic," conflating structural with dynamic inequalities. This is because these slopes are estimated while conditioning on the class destination linear component, and, as such, represent heterogeneous origin-mobility comparisons within a given class destination. In fact, it is only under very specific circumstances that Equations C.2 and C.3 will give unbiased estimates of structural and dynamic inequalities, respectively. Specifically, the Diff-SI model will produce the correct estimate of Γ_1 only if Γ_2 happens to be zero, while the DI-Diff model will give the correct estimate of Γ_2 only if Γ_1 happens to be zero. That is, $\Gamma_1 - \Gamma_2 = \Gamma_1$ only if $\Gamma_2 = 0$, and $\Gamma_2 - \Gamma_1 = \Gamma_2$ only if $\Gamma_1 = 0$. Thus, for the purposes directly estimating structural and dynamic inequalities, the SDI model is strongly preferred over the Diff-SI and DI-Diff models.

In the following sections, we examine the properties of all six logically possible one-factor and two-factor class models for a given mobility table or, equivalently, data with class origin, destination, and mobility variables. These models are listed in Table C.1. For each one- or two-factor model, we use a corresponding three-factor model to clarify exactly what is being estimated. For example, as shown in the first row of Table C.1, to understand the properties of the marginal destination model, which is a one-factor model, we use the SDI model. Similarly, to clarify the estimates of the two-factor origin-destination model (i.e., Duncan's "square additive model"), we use the DI-Diff model. Note that we consider all of these models to be descriptive, such that these various models are different ways of summarizing aggregate-level variability on a mobility table without relying on information purely external to the data.

Table C.1: Comparison of Class Models for a Mobility Table

Reference Model	One-Factor Model	Two-Factor Model
SDI model (C.1)	marginal destination (C.6)	origin-mobility (C.14)
Diff-SI model (C.2)	marginal mobility (C.8)	origin-destination (C.10)
DI-Diff model (C.3)	marginal origin (C.4)	destination-mobility (C.12)

Notes: This table outlines the various one- and two-factor models analyzed based on a corresponding reference model using matrix algebra and the logic of omitted variable bias. For example, the parameters of the marginal destination model and origin-mobility model are interpreted using the SDI model. Equation numbers are in parentheses. Note that these models are all treated as descriptive, and are thus distinct from a conventional mobility effects model, which has the form of $f(O^*, D^*, M^*)$, where O^*, D^*, and M^* are unobserved causal factors proxied by O, D, and M.

2. Interpreting the Parameters of One-Factor Class Models

In this section, we outline the three logically possible one-factor models (based on either origin, destination, or mobility) that can be used to describe the main patterns on a mobility table. For each one-factor model, we outline the relationship between the model's parameters and those from a corresponding model that includes all three factors (see Equations C.1, C.2, and C.3). To avoid confusion with corresponding terms in the three-factor models outlined previously, we use asterisks to denote the parameters in the one-factor models. In general, among all three one-factor models, we recommend using only the marginal destination model, as the underlying slope estimated by this model can be straightforwardly interpreted as a weighted sum of the ST and SM slopes from the SDI model.

i. Marginal Origin Model

The first logically possible one-factor model is the origin class mobility model, which has the following form:

$$\mu_{rijk} = f(O) = \mu^* + \alpha_i^* + \epsilon_{ri}^*, \tag{C.4}$$

where μ^* is the intercept; α_j^* are parameters for class origin using sum-to-zero deviation (or "effect") coding; and ϵ_{ri}^* denotes individual-level error. Using the DI-Diff model as a reference (see Equation C.3), the class origin model outlined in Equation C.4 can be shown to be equivalent to the following:

$$\begin{aligned} \mu_{rijk} &= \underbrace{(\mu + \phi_\mu)}_{\mu^*} + \underbrace{(\alpha_M + \phi_{\alpha_M})(i - i^*) + (\widetilde{\alpha}_i + \phi_{\widetilde{\alpha}_i})}_{\alpha_i^*} + \underbrace{(\epsilon_{rijk} + \eta_{ijk} + \nu_{ijk})}_{\epsilon_{ri}^*} \quad \text{and} \\ \alpha_M &= \big(\Gamma_1\omega_{(j,i)} + (\Gamma_2 - \Gamma_1)\omega_{(k,i)}\big), \end{aligned} \tag{C.5}$$

where μ is the intercept; Γ_1 is the ST slope; $\Gamma_2 - \Gamma_1 = \gamma - \alpha$ is the intra-destination slope; α_M is the marginal origin slope; $\omega_{(j,i)}$ is the relationship between the destination linear component and the origin linear component conditional on the intercept and origin nonlinear components; $\omega_{(k,i)}$ is the relationship between the mobility linear component and the origin linear component conditional on the intercept and origin nonlinear components;[54] $\widetilde{\alpha}_i$ is the ith origin nonlinearity; ϕ_μ, ϕ_{α_M} and $\phi_{\widetilde{\alpha}_i}$ are bias terms for the intercept, marginal origin slope, and the ith origin nonlinearity; ϵ_{rijk} is individual-level error; η_{ijk} denotes unique cell-specific heterogeneity; ν_{ijk} denotes additional heterogeneity attributable to class destination and mobility. The terms in brackets below Equation

[54]Those in the upper class can only be downwardly mobile or stay the same, while those in the lower class can only be upwardly mobile or stay the same. Accordingly, the relationship between the mobility and origin linear components is in general negative, and thus one can write $(\Gamma_2 - \Gamma_1)(-\omega_{(k,i)}) = (\Gamma_1 - \Gamma_2)\omega_{(k,i)}$ in Equation C.5.

C.5 denote the corresponding parameters from the marginal origin model presented in Equation C.4.

Several points are worth noting regarding Equation C.5. First, the marginal origin slope, or α_M, underlying Equation C.4 is a weighted sum of the ST and the intra-destination slope, with weights given by the relationships between destination and mobility, respectively, with origin. As noted in the main text, the intra-destination slope compares heterogeneous class-mobility groups, conflating structural with dynamic inequalities. For example, within a "middle" destination class, comparing a "low" class group with a "high" class group is simultaneously comparing a group that is upwardly mobile with another that is downwardly mobile. Because the marginal origin slope is in part a function of the intra-destination slope, we generally do not recommend using estimates from the marginal origin model.[55] Second, when the SM slope is zero (i.e., $\Gamma_2 = 0$), then the marginal origin slope will only be a function of the ST slope. In other words, in an absence of any observed social mobility, the marginal origin model will reflect overall structural inequalities. It is in this restricted sense that the marginal origin model could be used.[56] Third, the intercept, marginal origin slope, and origin nonlinearities will all have some degree of bias due to the exclusion of destination and mobility nonlinearities from Equation C.8.[57] Finally, the error term of the marginal origin model will reflect not just individual-level error, but also unique cell-specific heterogeneity as well as additional heterogeneity attributable to class destination and mobility.

ii. Marginal Destination Model

More commonly, researchers frequently a model that, while including other covariates, only includes class destination, omitting class origin and mobility (e.g., Goldthorpe 1999). The marginal class destination model, arguably the dominant model in sociology and demography, has the following general form:

$$\mu_{rijk} = f(D) = \mu^* + \beta_j^* + \epsilon_{rj}^*, \tag{C.6}$$

where μ^* is the intercept; β_j^* are parameters for class destination using sum-to-zero deviation (or "effect") coding; and ϵ_{rij}^* denotes individual-level error. Using the SDI model as a reference (see Equation C.1), the class destination model outlined in Equation C.6 can be shown to be equivalent to the following:

$$\mu_{rijk} = \underbrace{(\mu + \xi_\mu)}_{\mu^*} + \underbrace{\left(\beta_M + \xi_{\beta_M}\right)(j - j^*) + (\widetilde{\beta}_j + \xi_{\widetilde{\beta}_j})}_{\beta_j^*} + \underbrace{(\epsilon_{rijk} + \eta_{ijk} + \nu_{ijk})}_{\epsilon_{rj}^*} \quad \text{and}$$
$$\beta_M = \left(\Gamma_1 \omega_{(i,j)} + \Gamma_2 \omega_{(k,j)}\right), \tag{C.7}$$

where μ is the intercept; $\Gamma_1 = \alpha + \beta$ is the ST slope and $\Gamma_2 = \gamma + \beta$ is the SM slope; β_M is the marginal destination slope; $\omega_{(i,j)}$ is the relationship between origin linear component and the destination linear component conditional on the intercept and the destination nonlinear components; $\omega_{(k,j)}$ is the relationship between the mobility linear component and the destination linear component conditional on the intercept and the destination nonlinear components; $\widetilde{\beta}_j$ is the jth destination nonlinearity; ξ_μ, ξ_{β_M} and $\xi_{\widetilde{\beta}_j}$ are bias terms for the intercept, marginal destination slope, and

[55] One is generally better off using a model of the form $f(O, M)$. For additional discussion on the merits of models of this general form, see the main text.

[56] Note that, with data organized by origin, destination, and mobility, one can test whether or not the SM slope is zero or not using, for example, the SDI model.

[57] However, because our goal is primarily to conduct a descriptive rather than causal analysis, bias is less of a concern than understanding what, exactly, is being described with a particular model.

jth destination nonlinearity; ϵ_{rijk} is individual-level error; η_{ijk} denotes unique cell-specific heterogeneity; ν_{ijk} denotes additional heterogeneity attributable to class origin and mobility. The terms in brackets below Equation C.7 denote the corresponding parameters from the marginal destination model presented in Equation C.6.

Several main points are particularly noteworthy regarding Equation C.7. First, the marginal destination slope is a weighted sum of the ST and SM slopes, where the weights are given by the relationships between class origin and mobility, respectively, with class destination. Intuitively, this reflects the fact that class destination is a function of both structural and mobility processes. Accordingly, as an overall index of social stratification, class destination is in general a useful and informative metric. Second, the intercept, marginal destination slope, and destination nonlinearities will all have some degree of bias due to the exclusion of origin and mobility nonlinearities from equation C.6. Finally, the error term of the marginal destination model will reflect not only individual-level error, but also unique cell-specific heterogeneity on the mobility table, as well as additional heterogeneity attributable to class origin and mobility.

iii. Marginal Mobility Model

The third possible one-factor class model is the marginal class mobility model (e.g., see Chen et al. 2022), which has the following form:

$$\mu_{rijk} = f(M) = \mu^* + \gamma_k^* + \epsilon_{rk}^*, \tag{C.8}$$

where μ^* is the intercept; γ_j^* are parameters for class mobility using sum-to-zero deviation (or "effect") coding; and ϵ_{rik}^* denotes individual-level error. Using the Diff-SI model as a reference (see Equation C.2), the mobility model outlined in Equation C.8 can be shown to be equivalent to the following:

$$\begin{aligned} \mu_{rijk} &= \underbrace{(\mu + \psi_\mu)}_{\mu^*} + \underbrace{\big(\gamma_M + \psi_{\gamma_M}\big)(k - k^*) + (\widetilde{\gamma}_j + \psi_{\widetilde{\gamma}_k})}_{\gamma_k^*} + \underbrace{(\epsilon_{rijk} + \eta_{ijk} + \nu_{ijk})}_{\epsilon_{rk}^*} \quad \text{and} \\ \gamma_M &= \big((\Gamma_1 - \Gamma_2)\omega_{(i,k)} + \Gamma_2\omega_{(j,k)}\big), \end{aligned} \tag{C.9}$$

where μ is the intercept; $\Gamma_1 - \Gamma_2 = \alpha - \gamma$ is the intra-destination slope and $\Gamma_2 = \gamma + \beta$ is the SM slope; γ_M is the marginal mobility slope; $\omega_{(i,k)}$ is the relationship between origin linear component and the mobility linear component conditional on the intercept and mobility nonlinear components;[58] $\omega_{(j,k)}$ is the relationship between the destination linear component and the mobility linear component conditional on the intercept and mobility nonlinear components; $\widetilde{\gamma}_k$ is the kth mobility nonlinearity; ψ_μ, ψ_{γ_M} and $\psi_{\widetilde{\gamma}_k}$ are bias terms for the intercept, marginal mobility slope, and kth mobility nonlinearity; ϵ_{rijk} is individual-level error; η_{ijk} denotes unique cell-specific heterogeneity; ν_{ijk} denotes additional heterogeneity attributable to class origin and destination. The terms in brackets below Equation C.9 denote the corresponding parameters from the marginal destination model presented in Equation C.8.

As with the class destination model, several points are worth noting regarding Equation C.9. First, the marginal mobility slope, or γ_M, underlying Equation C.8 is a weighted sum of the intra-destination slope and the SM slope, with weights given by the relationships between origin and destination, respectively, with mobility. Again, as noted when discussing the marginal origin model, the intra-destination slope compares heterogeneous class-mobility groups, conflating structural with dynamic inequalities. Similarly, because the marginal mobility slope is a function of the intra-destination slope, we generally do not recommend using estimates from the marginal mobility

[58] Similar to the corresponding weight in Equation C.5, the relationship between the origin and mobility linear components is generally negative, and thus one can write $(\Gamma_1 - \Gamma_2)(-\omega_{(i,k)}) = (\Gamma_2 - \Gamma_1)\omega_{(i,k)}$ in Equation C.9.

model. Second, when the ST slope is zero (i.e., $\Gamma_2 = 0$), then the marginal mobility slope will only be a function of the SM slope. In other words, in an absence of any overall total social structural differences, the marginal mobility model will reflect the observed mobility patterns. It is in this restricted sense that the marginal mobility model could be used.[59] Third, as with the marginal destination model, the intercept, marginal mobility slope, and mobility nonlinearities will all have some degree of bias due to the exclusion of origin and destination nonlinearities from Equation C.8. Lastly, the error term of the marginal mobility model will capture not just individual-level error, but also unique cell-specific heterogeneity as well as additional heterogeneity attributable to class origin and destination.

3. Interpreting the Parameters of Two-Factor Class Models

In this section, we clarify the interpretation of the parameters from all three logically possible two-factor models (origin-destination, destination-mobility, origin-mobility). As with the one-factor models, we present each two-factor model and discuss the relationship between each model's parameters and those of a corresponding model that includes all three factors (see Equations C.1, C.2, and C.3). Again, to avoid confusion with corresponding terms in the three-factor models outlined earlier, we use asterisks to denote parameters from two-factor models. It should be emphasized that we treat these models as descriptive, not causal. In general, our analyses suggest that among the two-factor models, an origin-mobility model is preferable to an origin-destination or destination-mobility model. This is because the underlying linear terms of the origin-mobility model are the ST and SM slopes, which estimate structural and dynamic inequalities, whereas the other two models generate estimates of within-destination differences that compare heterogeneous origin-mobility groups.

i. Origin-Destination Model

The origin-destination model, also known as the "square additive model" (Hope 1971, 1975), has the following general form (cf. Duncan 1966: 94-95):

$$\mu_{rijk} = f(O, D) = \mu^* + \alpha_i^* + \beta_j^* + \eta_{ij}^* + \epsilon_{rij}^*, \qquad \text{(C.10)}$$

where μ^* is the intercept; α_i^* and β_j^* are parameters for origin and destination using sum-to-zero deviation (or "effect") coding; η_{ij}^* denotes group-level heterogeneity terms;[60] and ϵ_{rij}^* denotes individual-level error. Using the Diff-SI model (see Equation C.2), the origin-destination model outlined in Equation C.10 can be shown to be equivalent to the following:

[59] Note that, with data organized by origin, destination, and mobility, one can test whether or not the ST slope is zero or not using, for example, the SDI model.

[60] Following Duncan (1966: 94-95), we will treat the group-level heterogeneity terms for all of the two-factor models as orthogonal (see also Ohtaki et al. 1990: 119). These terms can be easily calculated as group-level residuals relative to a fully-saturated model. With respect to the origin-destination model, as an alternative one can specify all possible pairs of interactions between origin and destination. If the data are balanced such that there are an equal number of individual-level observations in each origin-destination cell, then, using sum-to-zero deviation coding or orthogonal polynomial coding, the residuals will be equivalent to specifying a full set of origin-destination interactions. The reason for this is that in such a setting the columns for the origin-destination interactions will be orthogonal to the main origin and destination columns.

$$\mu_{rijk} = \underbrace{(\mu + \psi_\mu)}_{\mu^*} + \underbrace{\big((\Gamma_1 - \Gamma_2) + \psi_{(\Gamma_1-\Gamma_2)}\big)(i - i^*) + (\widetilde{\alpha}_i + \psi_{\widetilde{\alpha}_i})}_{\alpha_i^*}$$
$$+ \underbrace{(\Gamma_2 + \psi_{\Gamma_2})(j - j^*) + (\widetilde{\beta}_j + \psi_{\widetilde{\beta}_j})}_{\beta_j^*} + \underbrace{(\eta_{ijk} + \nu_{ijk})}_{\eta_{ij}^*} + \underbrace{\epsilon_{rijk}}_{\epsilon_{rij}^*}, \tag{C.11}$$

where μ is the intercept; $\Gamma_1 - \Gamma_2 = \alpha - \gamma$ is a slope of intra-destination differences; Γ_2 is the DI slope; $\widetilde{\alpha}_i$ is the ith origin nonlinearity; $\widetilde{\beta}_j$ is the jth destination nonlinearity; ψ_μ, $\psi_{(\Gamma_1-\Gamma_2)}$, ψ_{Γ_2}, $\psi_{\widetilde{\alpha}_i}$, and $\psi_{\widetilde{\beta}_j}$ are bias terms for the intercept, intra-destination slope, DI slope, ith origin nonlinearity, and jth destination nonlinearity; η_{ijk} denotes terms for unique cell-specific heterogeneity; ν_{ijk} denotes terms for unique mobility-attributed heterogeneity; and ϵ_{rijk} is individual-level error. The terms in brackets below Equation C.11 denote the corresponding parameters from the origin-destination model presented in Equation C.10.

Three main points stand out from Equation C.11. First, the intercept, origin, and destination parameters will all have some degree of bias due to the exclusion of the mobility nonlinear components from the origin-destination model.[61] Second, assuming that there is no bias due to the exclusion of the mobility nonlinearities, either because the mobility nonlinearities are zero or the mobility variables are unrelated to the variables for the intercept, origin, and destination terms (i.e., the included variables), then the underlying origin and destination slopes of the origin-destination model will equal those from the Diff-SI model. In other words, the origin-destination model will generate a slope of intra-destination differences. For this reason we do not generally recommend using the origin-destination model without extreme care in the interpretation of the origin parameters in Equation C.10. Finally, the group-level heterogeneity terms η_{ij}^* from the origin-destination model equal the sum of the unique cell-specific heterogeneity terms from the Diff-SI model, η_{ijk}, and the unique mobility-attributed heterogeneity terms ν_{ijk}. The mobility-attributed heterogeneity terms are simply the predicted values from the parameters for the mobility nonlinear components (i.e., the excluded variables) using that part of the mobility variables that is unassociated with the variables for the intercept, origin, and destination terms (i.e., the included variables).

ii. Destination-Mobility Model

The second logically possible two-factor model is the destination-mobility model, which has the following general form:

$$\mu_{rijk} = \mu^* + \beta_j^* + \gamma_k^* + \eta_{jk}^* + \epsilon_{rjk}^*, \tag{C.12}$$

where μ^* is the intercept; β_j^* and γ_k^* are parameters for destination and mobility using sum-to-zero deviation coding; η_{jk}^* denotes group-level heterogeneity terms; and ϵ_{rjk}^* is individual-level error. Using the DI-Diff model (see Equation C.3), the destination-mobility model outlined in Equation C.12 can be shown to be equal to the following:

[61] Note that the individual-level η_{rik}^* error term is unbiased. The reason for this is that the origin-destination model with the group-level heterogeneity terms is saturated, so the individual-level error will be the same as that from the Diff-SI model with unique heterogeneity terms, which is also saturated. Again, because our focus here is on descriptive rather than causal models, we are less concerned about the parameters being biased than that researchers have a clear idea of what is being estimated in the models.

$$\mu_{rijk} = \underbrace{(\mu + \phi_\mu)}_{\mu^*} + \underbrace{(\Gamma_1 + \phi_{\Gamma_1})(j - j^*) + (\widetilde{\beta}_j + \phi_{\widetilde{\beta}_j})}_{\beta_j^*}$$
$$+ \underbrace{\big((\Gamma_2 - \Gamma_1) + \phi_{(\Gamma_2 - \Gamma_1)}\big)(k - k^*) + (\widetilde{\gamma}_k + \phi_{\widetilde{\gamma}_k})}_{\gamma_k^*} + \underbrace{(\eta_{ijk} + \nu_{ijk})}_{\eta_{jk}^*} + \underbrace{\epsilon_{rijk}}_{\epsilon_{rjk}^*}, \quad \text{(C.13)}$$

where μ is the intercept; $\Gamma_1 = \alpha + \beta$ is the ST slope; $\Gamma_2 - \Gamma_1 = \gamma - \alpha$ is the intra-destination slope; $\widetilde{\beta}_j$ is the jth destination nonlinearity; $\widetilde{\gamma}_k$ is the kth mobility nonlinearity; ϕ_μ, ϕ_{Γ_1}, $\phi_{(\Gamma_2 - \Gamma_1)}$, $\phi_{\widetilde{\beta}_j}$, and $\phi_{\widetilde{\gamma}_k}$ are bias terms for the intercept, ST slope, intra-destination slope, jth destination nonlinearity, and kth mobility nonlinearity; η_{ijk} denotes terms for unique cell-specific heterogeneity; ν_{ijk} denotes terms for unique origin-attributed heterogeneity; and ϵ_{rijk} is individual-level error. The terms in brackets below Equation C.13 denote the corresponding parameters from the destination-mobility model displayed in Equation C.12.

As with the origin-destination model, there are three main conclusions that follow from Equation C.11. First, as indicated by the presence of the ϕ parameters, the intercept, destination, and mobility parameters will be biased because of the exclusion of the origin nonlinearities from the destination-mobility model. Second, assuming that excluding the mobility nonlinearities results in no bias, then the underlying destination and mobility slopes will equal those from the DI-Diff model. Lastly, the group-level heterogeneity terms η_{jk}^* equal the sum of the unique cell-specific heterogeneity terms from the DI-Diff model, η_{ijk}, and the unique origin-attributed heterogeneity terms ν_{ijk}. Similar to the origin-destination model, the origin-attributed heterogeneity terms are just the predicted values from the parameters for the origin nonlinear components (i.e., the excluded variables) using that part of the origin variables that is unrelated to the variables for the intercept, destination, and mobility terms (i.e., the included variables).

iii. Origin-Mobility Model

The remaining two-factor model is the origin-mobility model, which has the following general form:

$$\mu_{rijk} = \mu^* + \alpha_i^* + \gamma_k^* + \eta_{ik}^* + \epsilon_{rik}^*, \quad \text{(C.14)}$$

where μ^* is the intercept; α_i^* and γ_k^* are parameters for origin and mobility using sum-to-zero deviation coding; η_{ik}^* denotes group-level heterogeneity terms; and ϵ_{rik}^* is individual-level error. Using the SDI model (see Equation C.1), the origin-mobility model presented in Equation C.14 is shown to be equivalent to the following:

$$\mu_{rijk} = \underbrace{(\mu + \xi_\mu)}_{\mu^*} + \underbrace{(\Gamma_1 + \xi_{\Gamma_1})(i - i^*) + (\widetilde{\alpha}_i + \xi_{\widetilde{\alpha}_i})}_{\alpha_i^*}$$
$$+ \underbrace{\big(\Gamma_2 + \xi_{\Gamma_2}\big)(k - k^*) + (\widetilde{\gamma}_k + \xi_{\widetilde{\gamma}_k})}_{\gamma_k^*} + \underbrace{(\eta_{ijk} + \nu_{ijk})}_{\eta_{ik}^*} + \underbrace{\epsilon_{rijk}}_{\epsilon_{rik}^*}, \quad \text{(C.15)}$$

where μ is the intercept; $\Gamma_1 = \alpha + \beta$ is the ST slope; $\Gamma_2 = \gamma + \beta$ is the SM slope; $\widetilde{\alpha}_i$ is the ith origin nonlinearity; $\widetilde{\gamma}_k$ is the kth mobility nonlinearity; ξ_μ, ξ_{Γ_1}, ξ_{Γ_2}, $\xi_{\widetilde{\alpha}_i}$, and $\xi_{\widetilde{\gamma}_k}$ are bias terms for the intercept, ST slope, SM slope, ith origin nonlinearity, and kth mobility nonlinearity; η_{ijk} denotes terms for unique cell-specific heterogeneity; ν_{ijk} denotes terms for unique destination-attributed heterogeneity; and ϵ_{rijk} is individual-level error. The terms in brackets below Equation C.15 refer to the corresponding parameters from the origin-mobility model shown in Equation C.14.

As with the other two-factor models, there are three main takeaways from Equation C.15. First, the intercept, origin, and mobility parameters will be biased because of the exclusion of the origin nonlinearities from the origin-mobility model. Second, assuming that excluding the destination nonlinearities produces no bias, then the underlying origin and mobility slopes will equal those from the SDI model. Finally, the group-level heterogeneity terms η^*_{ik} equal the sum of the unique cell-specific heterogeneity terms from the SDI model, η_{ijk}, and the unique destination-attributed heterogeneity terms ν_{ijk}. The destination-attributed heterogeneity terms are, like those for the other two-factor models, simply the predicted values from the parameters for the destination nonlinearities (i.e., the excluded variables) using that part of the destination variables that is unrelated to the variables for the intercept, origin, and mobility terms (i.e., the included variables).

4. Derivation of Relationships

In this section we show how the relationships outlined above can be derived using matrix algebra and the logic of omitted variable bias. We first present the derivation for the one-factor formulas using the one-factor destination model as an example. Next, we show the derivation for the two-factor models using the origin-destination model (i.e., Duncan's "square additive model").

i. Derivation for One-Factor Models

To show the derivation for the one-factor models, we using the class destination model, but similar calculations can be applied to derive the one-factor origin and mobility models. Suppose we fit the one-factor destination model (Equation C.6) on an individual-level data set indexed by origin, destination, and mobility. To reveal the underlying structure of the model, it is useful to express Equation C.6 as a linearized destination model, which decomposes each deviation from the overall mean into its constitutive linear and nonlinear components:

$$\mu_{rijk} = f(D) = \mu^* + \beta^*(j - j^*) + \widetilde{\beta}^*_j + \epsilon^*_{rj}, \tag{C.16}$$

where the parameters are the same as in Equation C.6 except β^* denotes the destination slope and $\widetilde{\beta}^*_j$ the jth destination deviation from the overall mean. Because only the coding scheme differs between Equation C.16 and Equation C.6, we will refer to them interchangeably as a class destination model in the following discussion.

Let $\mathbf{y}$ denote an $R \times 1$ column vector of outcome values (e.g., means), $\mathbf{1}$ an $R \times 1$ column vector of 1's, $\mathbf{d}_L$ an $R \times 1$ column vector of the class destination linear component, and $\widetilde{\mathbf{D}}$ an $R \times (J-2)$ matrix of orthogonal destination polynomials with no linear component. Using matrix notation, the marginal destination model in Equation C.16 can be expressed as follows:

$$\mathbf{y} = \mathbf{1}\mu^* + \mathbf{d}_L\beta^*_L + \widetilde{\mathbf{D}}\widetilde{\boldsymbol{\beta}}^* + \boldsymbol{\epsilon}^*. \tag{C.17}$$

where μ^* is again the intercept, β^*_M is the estimated marginal destination slope, $\widetilde{\boldsymbol{\beta}}^*$ is a $(J-2) \times 1$ column vector of nonlinear destination parameters, and $\boldsymbol{\epsilon}^*$ is an $R \times 1$ column vector of individual-level error terms.

For the purposes of comparison, note that the SDI model (see Equation C.1) can be specified in matrix form as:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{O}\boldsymbol{\alpha} + \mathbf{M}\boldsymbol{\gamma} + \widetilde{\mathbf{D}}\tilde{\boldsymbol{\beta}} + \boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{C.18}$$

where μ is the intercept, $\mathbf{O}$ is an $R \times (I-1)$ matrix of orthogonal origin polynomials, $\boldsymbol{\alpha}$ is an $(I-1) \times 1$ column vector of linear and nonlinear origin parameters, $\mathbf{M}$ is an $R \times (K-1)$ matrix of orthogonal mobility polynomials, $\boldsymbol{\gamma}$ is a $(K-1) \times 1$ column vector of linear and nonlinear mobility parameters, $\widetilde{\mathbf{D}}$ is an $R \times (J-2)$ matrix of orthogonal destination polynomials without the linear component, $\widetilde{\boldsymbol{\beta}}$ is a $(J-2)\times 1$ column vector of nonlinear destination parameters, $\boldsymbol{\eta}$ is an $R\times 1$ column vector of cell-specific heterogeneity terms, and $\boldsymbol{\epsilon}$ is an $R\times 1$ column vector of individual-level error terms.

To clarify the interpretation of the marginal destination model, we need to specify an auxiliary equation that expresses the association between those variables included in the marginal destination model and those excluded from the marginal destination model but included in the SDI model. To do so, we first define a matrix $\mathbf{S}$ of dimension $J\times(I+K-2)$ as follows:

$$\mathbf{S} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}, \tag{C.19}$$

where $\mathbf{X} = [\mathbf{1}\ \mathbf{d}_L\ \widetilde{\mathbf{D}}]$ is an $R\times J$ matrix of 1's, the destination linear component, and higher-order orthogonal destination polynomials; and $\mathbf{Z} = [\mathbf{O},\mathbf{M}]$ is an $R\times(I+K-2)$ matrix of orthogonal origin and mobility polynomials. The matrix $\mathbf{S}$ is simply a collection of parameters representing relationships between those variables included in the marginal destination model ($\mathbf{X}$) and those variables excluded from the marginal destination model but included in the SDI model ($\mathbf{Z}$). Using $\mathbf{Z}$, $\mathbf{S}$, and $\mathbf{X}$, we can accordingly define an auxiliary equation compactly as $\mathbf{Z} = \mathbf{XS} + \mathbf{U}_Z$ or, equivalently:

$$[\mathbf{O},\mathbf{M}] = \mathbf{1}\mathbf{s}_\mu + \mathbf{d}_L\mathbf{s}_{d_L} + \widetilde{\mathbf{D}}\mathbf{S}_{\widetilde{D}} + \mathbf{U}_{[O,M]}, \tag{C.20}$$

where $\mathbf{s}_\mu$ is a $1\times(I+K-2)$ row vector of parameters, $\mathbf{s}_{d_L}$ is a $1\times(I+K-2)$ row vector of parameters, $\mathbf{S}_{\widetilde{D}}$ is a $(J-2)\times(I+K-2)$ matrix of parameters, and $\mathbf{U}_{[O,M]}$ is an $R\times(I+K-2)$ matrix of error terms representing that part of $\mathbf{O}$ and $\mathbf{M}$ unrelated to the variables included in the marginal destination model (i.e., the intercept, destination linear component, and higher-order destination polynomials).[62]

To clarify the meaning of the parameters of the marginal destination model (Equation C.17), we can simply substitute Equation C.20 into Equation C.17. This is easily accomplished by re-writing Equation C.18 as $\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\boldsymbol{\zeta} + \widetilde{\mathbf{D}}\widetilde{\boldsymbol{\beta}} + \boldsymbol{\eta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\zeta}$ is an $(I+K-2)\times 1$ column vector of origin and mobility parameters such that:

$$\boldsymbol{\zeta} = \begin{pmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\gamma}\end{pmatrix}$$

We then just plug in $\boldsymbol{Z} = \mathbf{1}\mathbf{s}_\mu + \mathbf{d}_L\mathbf{s}_{d_L} + \widetilde{\mathbf{D}}\mathbf{S}_{\widetilde{D}} + \mathbf{U}_Z$ into this equation. After rearranging terms, we obtain the following:

$$\mathbf{y} = \mathbf{1}\underbrace{(\mu + \mathbf{s}_\mu\boldsymbol{\zeta})}_{\mu^*} + \mathbf{d}_L\underbrace{(\mathbf{s}_{d_L}\boldsymbol{\zeta})}_{\beta_L^*} + \widetilde{\mathbf{D}}\underbrace{(\widetilde{\boldsymbol{\beta}} + \mathbf{S}_{\widetilde{D}}\boldsymbol{\zeta})}_{\widetilde{\boldsymbol{\beta}}^*} + \underbrace{\boldsymbol{\epsilon} + \boldsymbol{\eta} + \mathbf{U}_Z\boldsymbol{\zeta}}_{\boldsymbol{\epsilon}^*}, \tag{C.21}$$

which reveals how the SDI model is related to the marginal destination model. Several points are worth emphasizing. First, the marginal destination term is a weighted sum of the ST and SM slopes (which are contained in $\boldsymbol{\zeta}$), with weights given by the relationships between the destination linear component and the origin and mobility linear components (which are contained in the row vector $\mathbf{s}_{d_L}$). Second, the parameters from the marginal destination model will all have some degree of bias due to the exclusion of the origin and mobility components. Depending on the structure of the data, the origin and mobility polynomials in $\mathbf{Z}$ will be more or less associated with the set of included variables, namely, the vector $\mathbf{1}$, destination linear component $\mathbf{d}_L$, and higher-order destination polynomials $\widetilde{\mathbf{D}}$.[63] If these relationships are strong, then the bias will be large, and the parameter

[62] Note that $\mathbf{s}_\mu$ is simply the first row of $\mathbf{S}$, $\mathbf{s}_{d_L}$ is the second row, and $\mathbf{S}_{\widetilde{D}}$ is rows 3 to J of $\mathbf{S}$.

[63] However, note that, because the marginal destination slope is defined by the weighted sum of the ST and SM slopes, the bias for the marginal destination slope is a function of only the relationship between the included variables and the

estimates for the intercept and destination nonlinear terms from the marginal destination model and the SDI models will differ, possibly quite substantially. By contrast, if these relationships are weak, then the bias will be relatively small, such that the intercept and destination nonlinear terms of the marginal destination model will be approximately equal to those from the SDI model. Lastly, the individual-level error term of the marginal destination model can ($\boldsymbol{\epsilon}^*$) be interpreted as the sum of individual-level error term from the SDI model, the unique cell-specific heterogeneity terms $\boldsymbol{\eta}$, and the column vector of origin and mobility parameters $\boldsymbol{\zeta}$, the latter of which are weighted by $\mathbf{U}_Z$, or that part of the excluded variables (i.e., the orthogonal origin and mobility polynomials) unrelated to the variables included in the marginal destination model.

ii. Derivation for Two-Factor Models

We illustrate the derivation for the two-factor models using the origin-destination model, but similar calculations can be applied to the destination-mobility and origin-mobility models. Suppose we fit the origin-destination model (Equation C.10) on an individual-level data set indexed by origin, destination, and mobility. To reveal the underlying structure of the model, it is useful to express Equation C.10 as a linearized origin-destination model with group-level heterogeneity terms, which decomposes each deviation from the overall mean into its constitutive linear and nonlinear components:

$$\mu_{rijk} = \mu^* + \alpha^*(i - i^*) + \widetilde{\alpha}_i^* + \beta^*(j - j^*) + \widetilde{\beta}_j^* + \eta_{ij}^* + \epsilon_{rij}^*, \quad \text{(C.22)}$$

where the parameters are the same as in Equation C.10 except α^* denotes the origin slope, $\widetilde{\alpha}_i^*$ the ith origin deviation from the overall mean, β^* the destination slope, and $\widetilde{\beta}_j^*$ the jth destination deviation from the overall mean. Because Equation C.22 is the same as that in Equation C.10, but with a different coding scheme, we will refer to them interchangeably as an origin-destination model in the discussion that follows.

Let $\mathbf{y}$ denote an $R \times 1$ column vector of outcome values (e.g., means), $\mathbf{1}$ an $R \times 1$ column vector of 1's, $\mathbf{O}$ an $R \times (I-1)$ matrix of orthogonal origin polynomials, and $\mathbf{D}$ an $R \times (J-1)$ matrix of orthogonal destination polynomials. Using matrix notation, the origin-destination model in Equation C.22 can be expressed as follows:

$$\mathbf{y} = \mathbf{1}\mu^* + \mathbf{O}\boldsymbol{\alpha}^* + \mathbf{D}\boldsymbol{\beta}^* + \boldsymbol{\eta}^* + \boldsymbol{\epsilon}^*. \quad \text{(C.23)}$$

where μ^* is again the intercept, $\boldsymbol{\alpha}^*$ is an $(I-1) \times 1$ column vector of linear and nonlinear origin parameters, $\boldsymbol{\beta}^*$ is a $(J-1) \times 1$ column vector of linear and nonlinear destination parameters, $\boldsymbol{\eta}^*$ is an $R \times 1$ column vector of group-level heterogeneity parameters, and $\boldsymbol{\epsilon}^*$ is an $R \times 1$ column vector of individual-level error terms.

For the purposes of comparison, note that the DI-Diff model (see Equation C.3) can be specified in matrix form as:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{O}\boldsymbol{\alpha} + \mathbf{D}\boldsymbol{\beta} + \widetilde{\mathbf{M}}\widetilde{\boldsymbol{\gamma}} + \boldsymbol{\eta} + \boldsymbol{\epsilon}, \quad \text{(C.24)}$$

where μ is the intercept, $\boldsymbol{\alpha}$ is an $(I-1) \times 1$ column vector of linear and nonlinear origin parameters, $\boldsymbol{\beta}$ is a $(J-1) \times 1$ column vector of linear and nonlinear destination parameters, $\widetilde{\mathbf{M}}$ an $R \times (K-2)$ matrix of orthogonal mobility polynomials with no linear component, $\widetilde{\boldsymbol{\gamma}}$ is a $(K-2) \times 1$ column vector of nonlinear mobility parameters, $\boldsymbol{\eta}$ is an $R \times 1$ column vector of cell-specific heterogeneity terms, and $\boldsymbol{\epsilon}$ is an $R \times 1$ column vector of individual-level error terms.

Similar to the calculations for the marginal destination model in the previous section, to interpret the meaning of the parameters of the origin-destination model, we need to specify an auxiliary equation that expresses the association between those variables included in the origin-destination

origin and mobility polynomials without the linear component.

model and those excluded from the origin-destination model but included in the DI-Diff model. As before, we can define a matrix $\mathbf{S}$ of dimension $(I+J-1)\times(K-2)$ as $\mathbf{S} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}$, where $\mathbf{X} = [\mathbf{1}\ \mathbf{O}\ \mathbf{D}]$ is an $R \times (I+J-1)$ matrix of 1's, orthogonal origin polynomials, and orthogonal destination polynomials; and $\mathbf{Z} = \widetilde{\mathbf{M}}$ is an $R \times (K-2)$ matrix of orthogonal class mobility polynomials with no linear component. Again, the matrix $\mathbf{S}$ is simply a collection of parameters representing relationships between those variables included in the origin-destination model ($\mathbf{X}$) and those variables excluded from the origin-destination model but included in the Diff-SI model ($\mathbf{Z}$). Using $\mathbf{Z}$, $\mathbf{S}$, and $\mathbf{X}$, we can accordingly define an auxiliary equation compactly as $\mathbf{Z} = \mathbf{XS}+\mathbf{U}_Z$ or, equivalently:

$$\widetilde{\mathbf{M}} = \mathbf{1s}_\mu + \mathbf{OS}_O + \mathbf{DS}_D + \mathbf{U}_{\widetilde{M}}, \tag{C.25}$$

where $\mathbf{s}_\mu$ is a $1\times(K-2)$ row vector of parameters, $\mathbf{S}_O$ is an $(I-1)\times(K-2)$ matrix of parameters, $\mathbf{S}_D$ is a $(J-1)\times(K-2)$ matrix of parameters, and $\mathbf{U}_{\widetilde{M}}$ is an $R\times(K-2)$ matrix of error terms representing that part of $\widetilde{\mathbf{M}}$ unrelated to the variables included in the origin-destination model.[64]

To clarify the meaning of the parameters of the origin-destination model (Equation C.22), we can simply substitute Equation C.25 into Equation C.23. After substituting and rearranging terms, we obtain the following equation:

$$\mathbf{y} = \mathbf{1}\underbrace{(\mu + \mathbf{s}_\mu\tilde{\boldsymbol{\gamma}})}_{\mu^*} + \mathbf{O}\underbrace{(\boldsymbol{\alpha} + \mathbf{S}_O\tilde{\boldsymbol{\gamma}})}_{\boldsymbol{\alpha}^*} + \mathbf{D}\underbrace{(\boldsymbol{\beta} + \mathbf{S}_D\tilde{\boldsymbol{\gamma}})}_{\boldsymbol{\beta}^*} + \underbrace{(\boldsymbol{\eta} + \mathbf{U}_{\widetilde{M}}\tilde{\boldsymbol{\gamma}})}_{\boldsymbol{\eta}^*} + \underbrace{\boldsymbol{\epsilon}}_{\boldsymbol{\epsilon}^*}, \tag{C.26}$$

which reveals how the Diff-SI model is related to the origin-destination model. As noted previously, the intercept, origin, and destination parameters from the origin-destination model will all have some degree of bias due to the exclusion of the mobility nonlinear components. Depending on the structure of the data, the orthogonal mobility polynomials in $\widetilde{\mathbf{M}}$ will be more or less related to the vector $\mathbf{1}$, orthogonal origin polynomials $\mathbf{O}$, and orthogonal destination polynomials $\mathbf{D}$. If these relationships are strong, then the bias will be large, and the parameter estimates from the origin-destination and the Diff-SI models will differ, possibly quite substantially. By contrast, if these relationships are weak, then the bias will be relatively small, such that the intercept, origin, and destination parameters of the origin-destination model will be approximately equal to those from the Diff-SI model. Similarly, the vector of group-level heterogeneity terms $\boldsymbol{\eta}^*$ in the origin-destination model, which can be interpreted as a restricted set of origin-destination interactions, is equal to a weighted sum of the cell-specific heterogeneity terms $\boldsymbol{\eta}$ and the mobility nonlinearities $\widetilde{\boldsymbol{\gamma}}$, the latter of which are weighted by $\mathbf{U}_{\widetilde{M}}$, or that part of the excluded variables (i.e., the orthogonal mobility polynomials without the mobility linear component dropped) unrelated to the variables included in the origin-destination model.

Equation C.26 additionally clarifies how the mobility nonlinearities can be viewed as a kind of "structured" interaction with respect to origin and destination.[65] We can show this relationship by taking the equation $\boldsymbol{\eta}^* = \boldsymbol{\eta} + \mathbf{U}_{\widetilde{M}}\widetilde{\boldsymbol{\gamma}}$ and solving for $\widehat{\boldsymbol{\gamma}}$, the mobility nonlinearities from the Diff-SI model. Because $\mathbf{U}_{\widetilde{M}}$ is non-square, it does not have a regular inverse. However, it has a Moore-Penrose generalized inverse that is equal to the left inverse of $\mathbf{U}_{\widetilde{M}}$. Solving for $\widetilde{\boldsymbol{\gamma}}$ gives us the following:

$$\begin{aligned}\widetilde{\boldsymbol{\gamma}} &= \mathbf{U}^+_{\widetilde{M}}(\boldsymbol{\eta}^* - \boldsymbol{\eta}) = (\mathbf{U}'_{\widetilde{M}}\mathbf{U}_{\widetilde{M}})^{-1}\mathbf{U}'_{\widetilde{M}}(\boldsymbol{\eta}^* - \boldsymbol{\eta}) \\ &= (\mathbf{U}'_{\widetilde{M}}\mathbf{U}_{\widetilde{M}})^{-1}\mathbf{U}'_{\widetilde{M}}\boldsymbol{\eta}^* \end{aligned} \tag{C.27}$$

[64] Note that $\mathbf{s}_\mu$ is simply the first row of $\mathbf{S}$, $\mathbf{S}_O$ is rows 2 to I of $\mathbf{S}$, and $\mathbf{S}_D$ is rows $I+1$ to $I+J-1$ of $\mathbf{S}$.

[65] What this means in practice is that the SDI model outlined in the main text is, in fact, interactive in the data although it is additive in the parameters.

where the plus $+$ denotes a Moore-Penrose generalized inverse and $(\mathbf{U}'_{M}\mathbf{U}_{\widetilde{M}})^{-1}\mathbf{U}'_{\widetilde{M}}$ is the left inverse of $\mathbf{U}_{\widetilde{M}}$.[66] Equation C.27 reveals that the mobility nonlinearities are equal to a regression model predicting heterogeneity from origin-destination interactions using that part of the orthogonal mobility polynomials unrelated to the intercept, origin, and destination variables included in the origin-destination model.[67] Similar derivations as those outlined in this section can be conducted in an analogous way for the destination-mobility and origin-mobility models.[68]

5. Higher-Order Interactions on a Mobility Table

So far little has been stated about the structure of the higher-order interactions beyond that for origin, destination, and mobility (i.e., the η terms in the various models discussed above). In this section we discuss higher-order interactions on mobility tables. As we illustrate, only a limited number of interactions can be included beyond the main set of parameters for origin, destination, and mobility. This reflects the fact that, descriptively, including all three main parameters means there is already a structured interaction captured by a three-factor model.

To illustrative the limited number of interactions that included, suppose we use orthogonal polynomial contrasts so that we have $I-2$, $J-2$, and $K-2$ columns for the higher-order terms of origin, destination, and mobility, respectively. Then the SDI model[69] can be represented as:

$$Y_{rijk} = \mu + (\alpha+\beta)o_L + (\gamma+\beta)m_L + \sum_{i=2}^{I-1}\alpha^i o_i + \sum_{j=2}^{J-1}\beta^j d_j + \sum_{k=2}^{K-1}\gamma^k m_k + \eta_{ijk} + \xi_{rijk}, \tag{C.28}$$

where η_{ijk} are cell-specific heterogeneity terms and ξ_{rijk} are individual-level errors. We treat these as residual (or orthogonal) to the main parameters in the model in the discussions above (see also Duncan 1966).

However, an alternative representation of the η_{ijk} terms is to specify them as higher-order interactions. However, because of the linear dependency among the variables, only a restricted set of interactions can be included (for a similar point, see Mason and Fienberg 1985). Specifically, given an origin-destination mobility table, one can specify the η_{ijk} terms as follows:

$$\eta_{ijk} = \sum_{j=2}^{J-1}\psi_{Lj}(o_L d_j) + \sum_{i=2}^{I-2}\sum_{j=2}^{J-1}\psi_{ij}(o_i d_j) \tag{C.29}$$

where the number of additional parameters above the baseline SDI model are $(I-2)(J-2)$. Note that these additional terms represent interactions of a smoothed origin curve with the destination

[66] Note that we can drop $\boldsymbol{\eta}$ from Equation C.27 because, by construction (see Equation C.1), it is unrelated to the orthogonal mobility polynomials such that $(\mathbf{U}'_{\widetilde{M}}\mathbf{U}_{\widetilde{M}})^{-1}\mathbf{U}'_{\widetilde{M}}\boldsymbol{\eta}$ will produce a $K-2$ column vector of zeros.

[67] The more the mobility variables are related to the heterogeneity from the origin-destination interactions, the larger in absolute value the size of the mobility nonlinearities. Note further that if the mobility variables are unrelated to the variables included in the origin-destination model, then $\mathbf{U}_{\widetilde{M}} = \widetilde{\mathbf{M}}$. Accordingly, Equation C.27 simplifies further to $(\widetilde{\mathbf{M}}'\widetilde{\mathbf{M}})^{-1}\widetilde{\mathbf{M}}'\boldsymbol{\eta}^*$. In other words, assuming the included and excluded variables are unrelated, we can simply take the mobility variables and use them to predict the heterogeneity from the origin-destination interactions to obtain the mobility nonlinearities. To the extent that the mobility variables are only weakly related to the intercept, origin, and destination variables in the origin-destination model, then this procedure will reproduce, within an error of approximation, the mobility nonlinearities from the Diff-SI model.

[68] However, note that, because there is inherent censoring on a mobility table with respect to mobility, one cannot include all mathematically pairwise interactions in a destination-mobility or origin-mobility model By contrast, although we have treated origin-destination interactions as cell-specific residuals, one could model them as all possible pairwise interactions on a mobility table.

[69] We use the SDI model for illustrative purposes here, but our results apply to any of the models discussed above.

nonlinear components. These interactions could be reversed so that the parameters represent interactions of a smoothed destination curve with a full set of origin nonlinear components.

To illustrate how the additional terms can be included in an L-APC Model, consider data from an origin-destination mobility table. Given a mobility table, one can incorporate additional heterogeneity using the particular set of origin-destination interactions outlined in Equation C.29. The ψ_{Lj} parameters are interactions terms between the origin linear component and the destination nonlinear components, while the ψ_{ij} parameters are interactions between the origin nonlinear components (except for the last origin nonlinear component) and the destination nonlinear components. This allows a smoothed origin curve to vary as a function of the destination nonlinear components.

To show what the full SDI model with higher-order origin-destination interactions would be, suppose there are $I = 5$ origin groups and $J = 5$ destination groups (and thus $K = I + J - 1 = 9$ mobility groups). Above the baseline SDI model, we can include $(I - 2)(J - 2) = 9$ additional parameters representing particular origin-destination interactions. Then the SDI model with fully specified higher-order origin-destination interactions is:

$$\begin{aligned}\mu_{rijk} = \mu + (\gamma+\beta)m_L + (\alpha+\beta)o_L + \alpha^2 o_2 + \dots \alpha^4 o_4 + \beta^2 d_2 + \dots \beta^4 d_4 + \gamma^2 m_2 + \dots \gamma^8 m_8 + \\ \beta_{L2}(o_L d_2) + \beta_{L3}(o_L d_3) + \beta_{L4}(o_L d_4) + \beta_{22}(o_2 d_2) + \beta_{23}(o_2 d_3) + \beta_{24}(o_2 d_4) + \\ \beta_{32}(o_3 d_2) + \beta_{33}(o_3 d_3) + \beta_{34}(o_3 d_4) + \xi_{rijk},\end{aligned} \tag{C.30}$$

where the additional terms represent intra-mobility heterogeneity, or heterogeneity with the diagonals of the mobility table. However, we caution against indiscriminately including these higher order terms directly in the main model without checking for multicollinearity. In general, including these additional interactions results in a full-rank design matrix (and thus the model is identified), but in practice these additional columns are highly collinear with the main columns of the baseline SDI model, resulting in highly unstable estimates.

Online Appendix D: Prior Contributions Using the DRM

As shown in Figure 1 of the main text, the DRM has been widely and increasingly used to estimate the effects of social mobility on a wide range of outcomes. In this appendix, we provide a schematic summary of the topics addressed in these contributions and the general direction of effects that have been estimated. The tables below highlight the findings from the key studies upon which Figure 1 rests.

Table D.1: Social and Cultural Attitudes

Author	Year	Outcome Measure(s)	Null	+	−	NC
Kulis	1987	Attitudes towards family	X	X		
Marshall and Firth	1999	Life satisfaction	X			
Tolsma et al.	2009	Ethnic attitudes	X	X		
Coulangeon	2013	Musical taste			X	
Daenekindt and Roose	2013	Aesthetic dispositions	X	X		
Daenekindt and Roose	2014	Musical taste		X		
Coulangeon	2015	Musical taste	X	X		
Turner	2017	Musical taste			X	
Sieben	2017	Child-rearing values	X			
Domański and Karpiński	2018	Culinary taste		X		
Schaeffer	2019	Discrimination perception			X	
Rotengruber and Tyszka	2021	Musical taste		X		
Creighton et al.	2022	Attitudes towards immigration	X	X		
Mijs et al.	2022	Meritocratic beliefs		X		

Notes: Positive effects are given by +, negative effects by −, and zero or near-zero effects by "Null." Those effects that are indeterminate are given as "NC" (Not Classifiable).

Table D.2: Health

Author	Year	Outcome Measure(s)	Null	+	−	NC
Monden and de Graaf	2013	Self-assessed health				X
Missinne et al.	2015	Health care use		X		
Yang	2016	Smoking and drinking			X	
van der Waal et al.	2017	Obesity	X			
Billingsley and Matysiak	2018	Fertility			X	
Dennison	2018	Drug use			X	
Präg and Richards	2018	Stress biomarkers	X			
Tarrence	2018	Self-rated health		X	X	
Steiber	2019	Subjective health satisfaction		X		
Gugushvili et al.	2020	Smoking and drinking		X		
Yang	2020	Smoking	X		X	
Veenstra and Vanzella-Yang	2021	Self-rated health		X		
Zelinska et al.	2021	Self-rated health	X			
Bulczak and Gugushvili	2022	Cardiometabolic risk			X	
Bulczak et al.	2022	Various health measures	X			
Graf et al.	2022	Biological aging		X		
Iveson et al.	2022	Self-rated health	X			
Kempel et al.	2022	Cardiometabolic risk	X			
Luo	2022	Fertility				X
Tarrence	2022	Self-rated health		X		

Notes: Positive effects are given by +, negative effects by −, and zero or near-zero effects by "Null." Those effects that are indeterminate are given as "NC" (Not Classifiable).

Table D.3: Political Outcomes

Author	Year	Outcome Measure(s)	Null	+	−	NC
De Graaf et al.	1990	Left/Right preferences		X		
Weakliem	1992	Left/Right voting		X		
Clifford and Heath	1993	Left/Right voting			X	
Nieuwbeerta and Graaf	1993	Left/Right voting			X	
Breen and Whelan	1994	Left/Right preferences	X		X	
Graaf et al.	1995	Left/Right preferences			X	
Nieuwbeerta, Paul	1995	Left/Right voting		X		
Nieuwbeerta et al.	2000	Left/Right voting	X			
Breen	2001	Left/Right preferences	X		X	
Paterson	2008	Left/Right attitudes				X
Daenekindt et al.	2018	Trust	X	X		
Yan	2019	Political Participation				X
Jaime-Castillo and Marqués-Perales	2019	Redistribution preferences	X			
Kraus and Daenekindt	2022	Attitudes towards multiculturalism		X		
McNeil	2022	Left/Right attitudes	X	X		
Mcneil and Haberstroh	2022	Opinion on Brexit		X		
Wilson et al.	2022	Redistribution preferences			X	

Notes: Positive effects are given by +, negative effects by −, and zero or near-zero effects by "Null." Those effects that are indeterminate are given as "NC" (Not Classifiable).

Table D.4: Psychological Wellbeing

Author	Year	Outcome Measure(s)	Null	+	−	NC
Houle and Martin	2011	Psychological distress	X			
Zang & de Graaf	2016	Happiness	X			
Daenekindt	2017	Dissociation	X	X		
Schuck and Steiber	2018	Subjective well-being		X		
Dhoore et al.	2019	Life-satisfaction, depression	X			
Gugushvili et al.	2019	Depression		X		
Jasper et al.	2019	Life-satisfaction	X			
Schuck	2018	Subjective well-being	X		X	
Zhao and Li	2019	Subjective well-being		X		
Engzell and Ichou	2020	subjective social status, perceived financial situation			X	
Gugushvili et al.	2021	Allostatic load		X		
Kaiser and Trinh	2021	Life-satisfaction		X		
Zelinska et al.	2021	Subjective well-being		X		
Kwon	2022	Subjective well-being	X			

Notes: Positive effects are given by +, negative effects by −, and zero or near-zero effects by "Null." Those effects that are indeterminate are given as "NC" (Not Classifiable).

Online Appendix E: Further Simulations of the Bias in the DRM

Here we present results from simulations with various values of the nonlinear effects for origin and destination. Note that in these simulations we assume that the proportionality constraint holds for the nonlinear origin and destination effects, but not necessarily for their respective linear effects. Regardless of the nonlinear effects, in each simulation we assume that the origin and destination linear effects are of the same magnitude and direction. We present the nonlinear effects in terms of squared components, but similar results could be obtained using alternative representations of the nonlinearities (e.g., deviations from the linear effects). To clarify the sensitivity of the DRM to the nonlinear effects, in all simulations we keep the true origin, destination, and mobility linear effects fixed at 0.250. The top half of Table E.1 displays the results across different values of the quadratic origin parameter, while the bottom shows the results across different values of the quadratic destination parameter. As with the previous tables of simulations, the shaded rows denote those specific simulations in which the data generating parameters happen to be recovered by the DRM.

Table E.1: Sensitivity of Estimated Mobility Linear Effect
to Values of True Origin and Destination Nonlinear Effects

	True DGP			DRM Estimates			Bias of $\widehat{\gamma}$				
	α^2	β^2	γ	$\widehat{\alpha}$	$\widehat{\beta}$	$\widehat{\gamma}$	$(\beta)(\widehat{w}_{\text{o}})$	$-$	$(\alpha)(\widehat{w}_{\text{d}})$	$=$	$\widehat{\gamma}-\gamma$
Varying α^2:	0.000	0.050	0.250	0.000	0.500	0.000	$(0.250)(\frac{0.050}{0.050})$	$-$	$(0.250)(\frac{0.000}{0.050})$	$=$	0.250
	0.050	0.050	0.250	0.250	0.250	0.250	$(0.250)(\frac{0.050}{0.100})$	$-$	$(0.250)(\frac{0.050}{0.100})$	$=$	0.000
	0.100	0.050	0.250	0.333	0.167	0.333	$(0.250)(\frac{0.100}{0.150})$	$-$	$(0.250)(\frac{0.050}{0.150})$	$=$	0.083
	0.250	0.050	0.250	0.417	0.083	0.417	$(0.250)(\frac{0.250}{0.300})$	$-$	$(0.250)(\frac{0.050}{0.300})$	$=$	0.167
	0.500	0.050	0.250	0.455	0.046	0.455	$(0.250)(\frac{0.500}{0.550})$	$-$	$(0.250)(\frac{0.050}{0.550})$	$=$	0.205
	1.000	0.050	0.250	0.476	0.024	0.476	$(0.250)(\frac{1.000}{1.050})$	$-$	$(0.250)(\frac{0.050}{1.050})$	$=$	0.226
Varying β^2:	0.050	0.000	0.250	0.500	0.000	0.500	$(0.250)(\frac{0.050}{0.050})$	$-$	$(0.250)(\frac{0.000}{0.050})$	$=$	0.250
	0.050	0.050	0.250	0.250	0.250	0.250	$(0.250)(\frac{0.050}{0.100})$	$-$	$(0.250)(\frac{0.050}{0.100})$	$=$	0.000
	0.050	0.100	0.250	0.167	0.333	0.167	$(0.250)(\frac{0.050}{0.150})$	$-$	$(0.250)(\frac{0.100}{0.150})$	$=$	-0.083
	0.050	0.250	0.250	0.083	0.417	0.083	$(0.250)(\frac{0.050}{0.300})$	$-$	$(0.250)(\frac{0.250}{0.300})$	$=$	-0.167
	0.050	0.500	0.250	0.046	0.455	0.046	$(0.250)(\frac{0.050}{0.550})$	$-$	$(0.250)(\frac{0.500}{0.550})$	$=$	-0.205
	0.050	1.000	0.250	0.024	0.476	0.024	$(0.250)(\frac{0.050}{1.050})$	$-$	$(0.250)(\frac{1.000}{1.050})$	$=$	-0.226

Notes: Number of origin and destination groups is set at $I = 5$ and $J = 5$, respectively, for all simulations. Sample size for each simulation is $R = 5,958$. Shaded rows indicate that the DRM recovers the true mobility linear effect. For all simulations the true origin and destination linear effects are fixed at 0.250 and the true nonlinear effects other than α^2 and β^2 are fixed at zero. For simplicity, and without loss of generality, we assume no random error. For all simulations we assume that the true origin and destination linear effects are $\alpha = \beta = 0.250$. The bias arises due to the fact that the underlying origin and destination linear effects do not obey the proportionality constraints of the estimated weights, which are a function of the underlying nonlinear effects. For example, in the top row the estimated weights for origin and destination are $\widehat{w}_o = 1.000$ and $\widehat{w}_d = 0.000$, respectively. Yet the actual weights needed to recover the true origin and destination effects are $w_o = 0.500$ and $w_d = 0.500$, respectively.

Online Appendix F: Additional Complications for Point-Identified or Partially-Identified Mobility Effects

So far little has been stated about the meaning of a mobility "effect." In this Appendix we outline the conceptualization of a mobility effect implied by the conventional mobility effects literature, outlining a number of challenges.

Graphical Causal Models

Figure F.1 presents the basic setup of a mobility effects model in terms of directed acyclic graphs (DAGs). As is common in the sociological literature, we treat these DAGs as the graphical embodiment of Pearl's (2009) nonparametric structural equation models (NPSEMs), using them to explicitly encode the underlying causal structure among the relevant variables. Panel (a) shows the graphical model for the causal effects of the origin (O^*), destination (D^*), and mobility (M^*) underlying factors on an outcome (Y) along with a background variable (X).[70] Filled circles denote observed variables while hollow circles denote unobserved variables. Note that, because the causal factors are unobserved, they are denoted with hollow circles. As well, for simplicity of presentation we have omitted idiosyncratic causes that affect the three underlying, unobserved factors (U_{O^*}, U_{D^*}, U_{M^*}) and the outcome (U_Y).

In a conventional mobility effects analysis, the observed variables O, D, and M, which have the natural relationship $D = O + M$, are substituted for the underlying causal variables O^*, D^*, and M^*. This scenario is shown in panel (b) of Figure F.1. The double lines indicate the linear dependency among the dimensions, such that $D := O + M$, where $:=$ means "is defined as." Note that these variables, unlike O^*, D^*, and M^*, are observed, as indicated by the solid points. As well, the observed variables O, D, and M are not affected by idiosyncratic causes, as they are deterministically related dimensions of the mobility table.

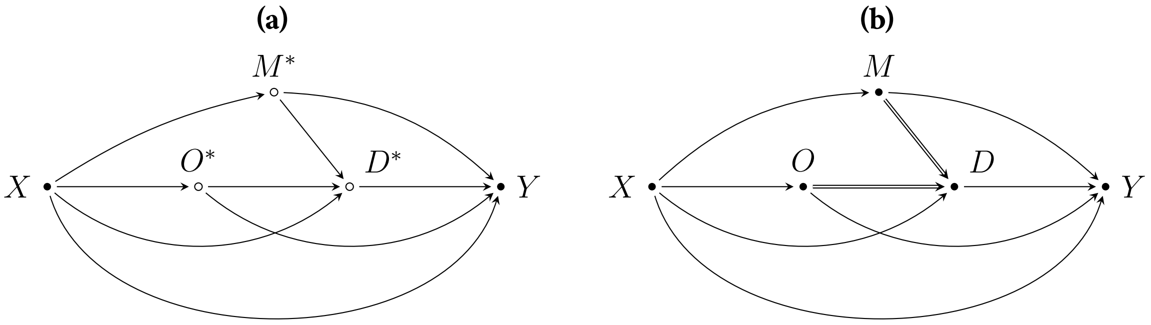
As we have discussed in previous sections, the scenario in Figure F.1(b) is highly problematic because of the linear dependence among origin, destination, and mobility. Even assuming that the underlying, unobserved factors are additive, with minimal interactions,[71] there is still not enough information to uniquely estimate all three causal effects. In operational terms, this means that we can only condition on two of the three variables (or, more precisely, their linear components). This is clearly a problem, as the graphical model in Figure F.1 shows that estimating just two of the three causal effects will lead to biased estimates.

For the sake of the present discussion, however, let us assume that we have somehow obtained the underlying causal factors O^*, D^*, and M^*. Suppose further that we are somehow able to obtain estimates that correspond to causal graph in Figure F.1 (a). Even in such an idealized scenario, serious problems remain in interpreting the estimates of these effects. We outline four such problems below. It is important to understand that these issues arise in all models of mobility effects, and add a further layer of complexity to identification and estimation.

[70] As outlined previously, in conventional models of mobility effects, origin, destination, and mobility are implicitly treated as surrogates for unobserved factors that actually generate an outcome of interest. Again, let O^*, D^*, and M^* denote underlying causal factors that are allowed to freely vary from each other such that $D^* \neq O^* + M^*$.

[71] It is important to note that saying that O^*, D^*, and M^* are additive is not the same as saying that O, D, and M are additive. The reason is that the unobserved factors lie on a three-dimensional tensor, while the observed dimensions lie on a two-dimensional mobility table. In the two-dimensional mobility table, the nonlinearities in any one of the dimensions will appear "interactive" with respect to the other two dimensions. For example, the class destination nonlinearities will appear at different mobility levels for different class origin groups.

Figure F.1: Graphical Models of Origin, Destination, and Mobility Effects



Notes: Panel **(a)** shows the graphical model for the causal effects of the origin (O^*), destination (D^*), and mobility (M^*) underlying, unobserved factors on an outcome (Y) along with a background variable (X). Filled circles denote observed variables while hollow circles denote unobserved variables. Idiosyncratic causes that affect the three underlying factors (U_{O^*}, U_{D^*}, U_{M^*}) and the outcome (U_Y) are omitted for simplicity of presentation. Panel **(b)** shows the graphical model with the observed origin (O), destination (D), and mobility (M) dimensions used as proxies for the for the underlying, unobserved causal factors. The double lines indicate the linear dependency among the dimensions, such that $D := O + M$, where $:=$ means "is defined as." Note that the observed variables O, D, and M are not affected by idiosyncratic causes.

Parallel-World Counterfactuals

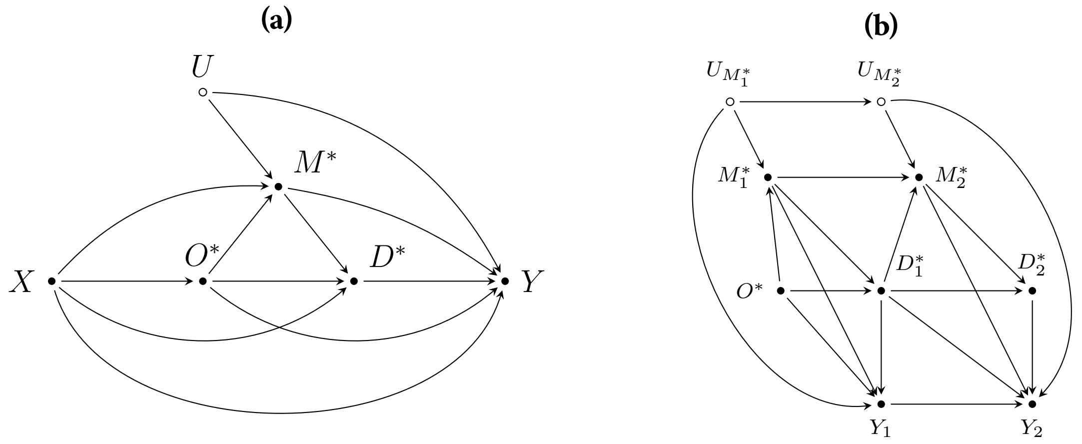
In the context of mobility effects models, counterfactuals can be defined in terms of intervening to fix (or set) some values of underlying origin, destination, and mobility factors for an individual. Let $Y^{o^* d^* m^*}$ be the outcome if we were to somehow intervene to fix the underlying factors of origin, destination, and mobility for a given individual. We could thus define, for a particular individual, a counterfactual outcome as $Y^{o^*=\text{low}, d^*=\text{high}, m^*=\text{down}}$, where low, high, and down refer to bundles of causal processes. Essentially, this counterfactual requires intervening to fix a set of causal factors for low class origin, high class destination, and downward mobility, thereby invoking, within the same individual, parallel worlds. In this case, the individual is fixed to values of the underlying mechanisms that imply upward mobility, for example, while also being fixed to mechanisms that imply downward mobility. This is not intrinsically problematic, as various estimated causal effects, such as natural direct and indirect effects, invoke parallel-world counterfactuals. It does mean, however, that there is no obvious way for a real-world intervention, such as a randomized experiment, to generate the expected value of this counterfactual for a given population.[72]

The Consistency Assumption

Suppose we believe that we have identified a causal effect for mobility. To endow this estimate with a formal counterfactual interpretation, one must invoke the assumption of consistency (Hernan 2016). Formally, given an individual respondent r is exposed to $M_r = m$, the counterfactual outcome Y_r^m is said to be "consistent" with the observed outcome Y_r if $Y_r^m = Y_r$. This assumption is violated if there are multiple ways to obtain a given level of exposure, thereby generating different counterfactuals. This assumption is particularly likely to be violated for composite variables that reflect multiple underlying features of the data. For example, the causal effect of cholesterol on health is "inconsistent" because the effect on health is very different depending on whether or

[72]Furthermore, as with the literature on natural direct and indirect effects, which has drifted toward so-called "interventional analogues" of mediator effects (Vansteeldant and Daniel 2017), it is unclear whether such parallel-world counterfactuals are actually of interest to applied researchers.

Figure F.2: Social Mobility as a Confounder



Notes: Panel (a) shows the underlying graphical model, which now includes a path between O^* and M^* as well as an additional unobserved causal variable U. It is assumed that the underlying causal variables O^*, D^*, and M^* are all observed. The causal effect of O^* is confounded due to M^*. However, standard adjustment techniques introduce potential bias because M^* is also a mediator on the causal path from O^* to Y. An additional complication is introduced by an unobserved confounder U between class mobility and the outcome. With U, the mobility variable M^* becomes a collider variable even after adjusting for X. Thus, adjusting for M^* using standard techniques will open a backdoor path between O^* and Y, thereby biasing the estimated causal effect of class origin on the outcome. Panel (b) shows the graphical model if there are multiple destination destination and mobility causal factors measured at subsequent time periods (D_1^*, M_1^*, D_2^*, M_2^*), with multiple outcomes (Y_1, Y_2) and multiple unobserved confounders ($U_{M_1^*}$, $U_{M_2^*}$). The same issues arise in this more complicated setting.

not one intervenes to raise cholesterol by increasing HDL ("good" cholesterol) versus LDL ("bad" cholesterol). The practical advice is to use variables that correspond to narrowly defined exposures, with correspondingly well-defined counterfactuals. This can be viewed as a subjective decision that depends on one's expertise and tolerance for ambiguity. For example, Rehkopf et al. (2016) outline ways in which neighborhoods, income, and education can each be understood to violate the consistency assumption, despite widespread agreement in sociology and related fields that these are all causal variables. For instance, they consider education as an example that violates the consistency assumption, inasmuch it has an effect on the outcome via, for example, "improvements in knowledge and cognitive skills, credentials that are valued on the labor market, status improvements, and changes to the individual's social network." (2016: 66). Unfortunately, measures of social class and class mobility are clearly a composite variable that reflects a variety of underlying mechanisms (e.g., Wright 2005) and thus leads to ill-defined, ambiguous counterfactuals. This suggests that the basic mobility effects approach, in which origin, destination, and mobility are proxies for omnibus underlying causal factors, should be replaced by a more targeted approach that focuses on specific, well-defined causal mechanisms.

Unobserved Confounding

To identify the causal effect of social mobility, we must assume that mobility is not confounded with the outcome. This implies conditioning on relevant background variables and avoiding conditioning on post-exposure confounders, which would block some of the effects of mobility. However, the assumption of no unobserved confounding is particularly thorny with respect to mobility effects models. This is closely related to the consistency assumption. Because these are omnibus factors representing multiple causal mechanisms and background variables, it is difficult to imagine which unobserved variables would confound the effect of mobility on a given outcome. Again, more narrowly defined exposures are helpful in figuring out what variables are potential confounders, but this will lead the researcher away from the analysis of mobility effects as they have been conventionally understood in the literature.

Mobility As A Confounder

Finally, a fourth problem with models of mobility effects concerns the fact that social mobility itself is a confounding variable. A careful inspection of Figure F.1 reveals that we have assumed that the underlying origin variable has no causal effect on the underlying mobility variables. This is unrealistic in practice. Presumably, the bundle of causal factors for class origin affects not only those for class destination, but also mobility. This scenario is illustrated in Figure F.2(a), which now includes a path between O^* and M^* as well as an additional unobserved causal variable U. To reiterate, following the assumptions of our present discussion, it is assumed that the underlying causal variables O^*, D^*, and M^* are all observed.

Suppose we want to identify the causal effect of D^*, which is now confounded by M^*. Assuming we have observed these underlying variables, standard adjustment techniques would introduce potential bias because M^* is also a mediator on the causal path from O^* to Y. An additional complication is introduced by an unobserved confounder U between class mobility and the outcome. With U, the mobility variable M^* becomes a collider variable. Thus, adjusting for M^* using standard techniques would open a backdoor path between O^* and Y, thereby biasing the estimated causal effect of class origin on the outcome. The same issues extend to more complicated settings, such as that shown in Figure F.2(b), with multiple class destination and mobility variables, as well as multiple confounders.[73]

[73] Assuming one has observed factors for origin, destination, and mobility, then various methods for time-varying confounding could be used to estimate the effects, such as structural nested models (Vansteeldant and Joffe 2014), marginal structural models (Robins et al. 2000), or residualized regression models (Wodtke and Xiang 2020). However, the problems of consistency, no unobserved confounding, and parallel counterfactuals would remain.

Online Appendix G: Additional Structural and Dynamic Inequalities

In the manuscript, we focus on three parametric expressions that can be derived from the Structural and Dynamic Inequality (SDI) Model, namely the *Social Structure Slope* and the *Social Mobility Slope*, *Social Structure Curves* and *Social Mobility Curves*, and *Comparative Mobility Curves*. Below, we first show the full list of expressions that can be derived from the SDI model and then illustrate the analytic use of one additional expression in the form of the *Marginal Class Destination Curve*.

Full list of expressions derived from the SDI model

Table G.1: Summarizing Structural and Dynamic Inequalities on a Mobility Table

General Terminology	Specific Summary	Mathematical Expression	
Structural Inequality	Social Structure Slope	$\Gamma_1(i - i^*)$	for all i
	Social Structure Curve	$\Gamma_1(i - i^*) + \widetilde{\alpha}_i$	for all i
	Social Structure Surface	$\Gamma_1(i - i^*) + \widetilde{\alpha}_i + \widetilde{\beta}_j$	for combinations of i, j
	Local Social Structure Curves	$\Gamma_1(i - i^*) + \widetilde{\alpha}_i + \widetilde{\beta}_{i+k-I}$	for all i in each mobility group k
Dynamic Inequality	Social Mobility Slope	$\Gamma_2(k - k^*)$	for all k
	Social Mobility Curve	$\Gamma_2(k - k^*) + \widetilde{\gamma}_k$	for all k
	Social Mobility Surface	$\Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_j$	for combinations of k, j
	Local Social Mobility Curves	$\Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{k+i-I}$	for all k in each origin group i
Structural & Dynamic Inequalities	Adjusted Marginal Destination Slope	$\Big(\Gamma_1\omega_{(i,j)} + \Gamma_2\omega_{(k,j)}\Big)(j - j^*)$	for all j
	Adjusted Marginal Destination Curve	$\Big(\Gamma_1\omega_{(i,j)} + \Gamma_2\omega_{(k,j)}\Big)(j - j^*) + \widetilde{\beta}_j$	for all j
	Overall Comparative Mobility Curve	$\phi_i + \Gamma_2(k - k^*) + \widetilde{\gamma}_k$	for all k in each origin group i
	Adjusted Comparative Mobility Curve	$\phi_i + \Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{i+k-I}$	for all k in each origin group i
	Unadjusted Comparative Mobility Curve	$\phi_i + \Gamma_2(k - k^*) + \widetilde{\gamma}_k + \widetilde{\beta}_{i+k-I} + \eta_{i[i+k-I]k}$	for all k in each origin group i

Notes: $\Gamma_1 = \alpha + \beta$ and $\Gamma_2 = \gamma + \beta$. The quantity ϕ_i is equal to $\Gamma_1(i - i^*) + \widetilde{\alpha}_i$, which is a single value for a given origin group i.

Adjusted Marginal Destination Curves

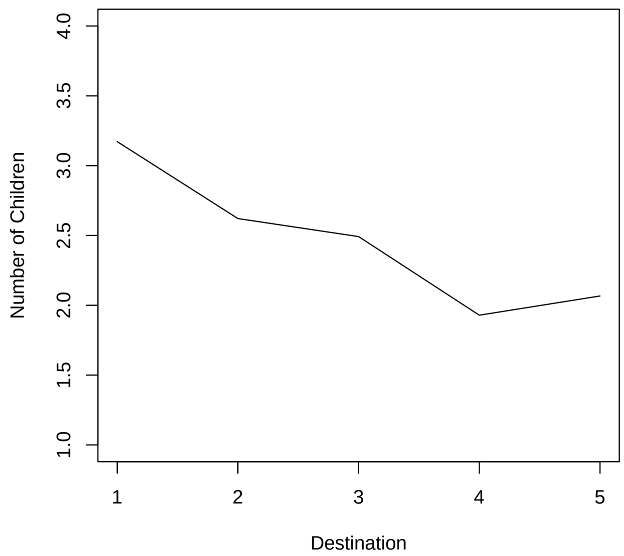
To illustrate the analytic use of just one additional expression from Table G.1, we now discuss the *Adjusted Marginal Destination Curve*. So far we have focused on examining the data through the lens of the SDI model, which is a general model of the form $Y = f(O, M) + \epsilon$. However, as we outlined in the previous section, it may be useful in some circumstances to also examine the data using a marginal class destination model, which has the general form of $Y = f(D) + \epsilon$, where again, without a loss of generality, ϵ is a normally distributed error term with a mean of zero. A particularly useful summary is what we call the *adjusted marginal destination curve*, which is equal to:

$$\beta_M(j - j^*) + \widetilde{\beta}_j = \Big(\Gamma_1\omega_{(i,j)} + \Gamma_2\omega_{(k,j)}\Big)(j - j^*) + \widetilde{\beta}_j \text{ for } j = 1, \ldots, J, \quad \text{(G.1)}$$

where $\omega_{(i,j)}$ is the relationship between the origin linear component and the destination linear component conditional on the origin, destination, and mobility nonlinearities while $\omega_{(k,j)}$ is the relationship between the mobility linear component and the destination linear component again conditional on the origin, destination, and mobility nonlinearities. This curve is equivalent to a simple class destination model (see Equation C.6 in Online Appendix C), but we have adjusted for the origin, destination, and mobility nonlinearities. Failing to adjust for the origin and mobility nonlinearities will introduce bias into the estimated overall class destination gap. However, a more practical reason for adjusting for the origin and mobility nonlinearities is that we can decompose the overall (linear) class destination gap into structural and dynamic components. This can thus answer crucial questions regarding the extent to which cross-destination differences are attributable to differences in the social structure versus social mobility.

The $\widehat{\Gamma}_1\widehat{\omega}_{(i,j)}$ term in Equation G.1 gives the contribution of social structure to the adjusted class destination slope, while $\widehat{\Gamma}_2\widehat{\omega}_{(k,j)}$ gives the contribution of social mobility. In general, social structure will contribute more to the class destination slope the greater the degree of correlation between class origin and destination, as well as the greater the relationship between the ST slope and the outcome. Similarly, in general, social mobility will contribute more to the class destination slope the greater the degree of correlation between class mobility and destination, as well as the stronger the relationship between the SM slope and the outcome. In a world with no differences in social mobility, the class destination gap will be driven entirely by the social structure; conversely, in a world without any structural inequality, the destination gap will be entirely a function of social mobility. In other words, the SDI model can be used to decompose any marginal class destination gap, i.e., the kind of social class gaps that are arguably the most common estimand in social stratification research, into distinct structural- versus mobility-based components.

Figure G.1: Marginal Class Destination Curve



Notes: Panel shows the adjusted marginal class destination curve, which is estimated conditional on the origin and mobility nonlinearities. The curve is a function of $\beta_M(j - j^*) + \widetilde{\beta}_j$ for all class destination groups j. Data are based on Sobel (1981).

Figure G.1 shows the adjusted marginal class destination curve for the fertility data. As can be seen, there is a general observed decline in fertility as one compares lower versus higher class destinations. The underlying adjusted marginal destination slope is $\widehat{\beta}_M = -0.290$, indicating a negative relationship between fertility and class destination. Using Equation G.1, we can decompose this overall slope into structural and social mobility components:

$$\widehat{\beta}_M = \left(\widehat{\Gamma}_1 \widehat{\omega}_{(i,j)} + \widehat{\Gamma}_2 \widehat{\omega}_{(k,j)}\right) = (-0.317)(0.740) + (-0.213)(0.260) = (-0.235) + (-0.055), \qquad \text{(G.2)}$$

the sum of which equals adjusted marginal destination slope, or -0.290. In this case, most of the class destination gap is a function of structural differences, reflecting both the relatively large ST slope as well as the strong relationship between the class destination linear component and class destination origin linear component.