Intergenerational social mobility in the United States: a multivariate analysis using Bayesian distributional regression

Alexander März[a], Nadja Klein[b], Thomas Kneib[c], Oliver Mußhoff[d]

[a] *Author for correspondence*
Department of Agricultural Economics, Georg-August-University Göttingen
Email: Alexander.Maerz@agr.uni-goettingen.de

[b] *Chair of Statistics, Georg-August-University Göttingen*

[c] *Chair of Statistics, Georg-August-University Göttingen*

[d] *Department of Agricultural Economics, Georg-August-University Göttingen*

Abstract

Ever since its establishment, equal access to opportunity and the prospect of upward social mobility have been central to the United States as the land of opportunity. Recent research, however, suggests that social mobility varies considerably by geographic location. Regarding the measurement of social mobility, measures of absolute and relative mobility have traditionally been analyzed independently from one another. Yet, since many Americans think of the American Dream in both absolute and relative terms, it appears much more adequate to analyze the chances of Americans to move beyond their parents in terms of higher incomes, while changing their relative position in the income distribution at the same time. Using data for the United States at the county and commuting zone levels, we extend the literature by analyzing the joint occurrence of measures of absolute and relative social mobility within a multivariate framework. In contrast to previous studies, we relax the assumption of independence between measures of absolute and relative mobility and study their interdependence by relating the coefficient of correlation between measures of social mobility to covariates within a regression setting. This allows us to identify the factors that govern the strength of their association and to analyze the pathways through which absolute and relative mobility are linked. We find that the coefficient of correlation between measures of social mobility is non-constant and varies considerably, both with respect to economic variables and across the United States. Moreover, we add to a recent strand of literature in economics and social science, and analyze the conditional variance of social mobility within a heteroscedastic regression framework. By modelling heteroscedasticity as a function of economic factors, we are able to study the distribution of social mobility in more detail and to identify candidate factors which govern the insecurity attached to participate in upward social mobility. Our results suggest a Great Gatsby Curve for the variance, so that higher rates of income inequality do not only reduce the overall level of social mobility, but also increase the risk of being excluded from upward social mobility. With our analysis, we are also able to graphically assess the ability of the factors, as discussed in the sociology and economics literature, to capture the spatial patterns of social mobility across the United States. The results indicate that although these factors contribute to a better understanding of its underlying mechanisms, they leave spatial heterogeneities of social mobility largely unexplained. *JEL Codes: C11, C14, H00, J00, R00, R12.*

Keywords: Conditional dependence, Equality of opportunity, Heteroscedastic regression, Intergenerational social mobility.

1. Introduction

The common belief that through ambition, hard work and talent one has a fair shot of making it from a humble beginning to great wealth is inseparably tied to the United States as the land of opportunity. Based on this idea, the United States has evolved as a society that attaches greater importance to providing everyone with equal access to opportunity, rather than to guaranteeing income equality (The Pew Charitable Trusts, 2009). This firm conviction is mirrored by the tolerance of many Americans towards economic inequality, as long as social mobility is high enough to provide everyone with an equal chance of climbing the income ladder, independent of one's family background or social standing (Haskins et al., 2008).

Regarding the measurement of intergenerational social mobility, a general distinction between measures of absolute and relative social mobility can be made (Fields and Ok, 1999). Measures of absolute mobility focus on the overall financial situation, generally measured in monetary units such as income or wealth, and assess whether the current generation is financially better off than the previous one on an absolute income basis. Therefore, a society is said to have high levels of absolute upward mobility, if the overall standard of living has increased from one generation to the next (Bengali and Daly, 2013). In contrast, measures of relative social mobility focus on the rank order and assess whether children are able to improve upon their rank within the income distribution relative to their parents (Bengali and Daly, 2013). Consequently, a fluid society with high rates of relative mobility provides everyone with an equal chance of moving along the income distribution, where family background should have little to no impact on the future income position of children (Narey, 2009).

Although the Unites States has traditionally been perceived as a society that is supposedly characterized by high levels of social mobility, existing research shows that the empirical verification is dependent upon whether the current situation of the United States as the land of opportunity is assessed through the perspective of either absolute or relative mobility. A recent overview provided by the Pew Charitable Trust shows that overall, 84% of today's adult children have a family income that exceeds that of their parents (Urahn et al., 2012). Comparing these numbers across different quintiles of the income distribution shows that 93% of the children raised in families with an income in the bottom fifth have a higher family income than their parents, with the corresponding percentages being 86%, 88%, 85% and 70% for the second, middle, fourth and the top fifth of the income distribution. Consequently, the American dream is doing reasonably well in absolute terms, as most Americans achieve a standard of living that is greater than that of their parents. However, when the state of social mobility is asssessed in terms of relative mobility, the United States is lagging behind internationally. While children born to parents in the middle of the distribution are more or less able to move freely within the income distribution as adults, 43% of today's adult children whose parents have incomes in the bottom fifth of the distribution will also have an income in the bottom fifth during adulthood. Likewise, among children born to parents in the top fifth, 40% find themselves in the top fifth of the income distribution as adults (Urahn et al., 2012). Hence, from a relative point of view, the picture of the United States as a socially mobile society is clouded, as the position of parents is still a strong predictor for the future income position of children, especially for those born to parents in the bottom and top fifth of the income distribution.[1] Moreover, a recent study by Chetty et al. (2014) reveals that there is substantial geographic variation in social mobility within the United States: in some parts of the United States, children in the bottom quintile have less than a 5% chance of reaching the top quintile, whereas in other areas, children in the bottom quintile have more than a 15% chance of reaching the top quintile.

What emerges from the preceding discussion on measures of social mobility is that the state of social mobility within a society can only be comprehensively assessed through a combined analysis of measures of absolute and relative mobility, since focusing only on one of the two measures fails to consider important aspects of the overall picture (Haskins et al., 2008). Although there is an increasing awareness of the multifaceted dimension of social mobility, existing studies studies have traditionally analyzed measures of absolute and relative mobility independently from each other. However, since the promise of the American Dream depends on a combination of both relative and absolute mobility, it appears much more adequate to analyze the chances of Americans to move beyond their parents in terms of higher incomes and in changing relative positions in the income distribution at the same time (Haskins et al., 2008). An initial step towards a joint modeling of absolute and relative mobility was taken by Haskins et al. (2008) and Urahn et al. (2012) who investigate the chances of children in contemporary America to move beyond their parents in absolute income levels, as well as in relative economic standing.[2] Using a new typology that combines the two measures of social mobility, Urahn et al. (2012) find that among all children in the United States, about 35% are upwardly mobile in the sense that they surpass their parents in both absolute and relative terms. However, the study reports that American children are also downwardly mobile, as 16% fall behind their parents in both dimensions of social mobility. The remaining children (49%) either stay in the same position of the income distribution as their parents or move down the income distribution in relative terms, despite having family incomes exceeding that of their parents.

Even though the studies of Haskins et al. (2008) and Urahn et al. (2012) make important contributions towards a more comprehensive analysis of social mobility, they do not provide insight into the factors that influence the joint occurrence of measures of absolute and relative social mobility. However, knowledge of the impact of how economic and social factors act on the mutual dependence between measures of absolute and social mobility is of importance since both types of social mobility matter and since both measures can reinforce the other: higher rates of relative mobility may reflect a more competitive labour market, which leads to higher rates of economic growth and, consequently, to higher rates of absolute mobility (Reeves and Sawhill, 2014). Consequently, from a policy-maker point of view, it is essential to better understand the interdependencies between the two in order to balance the need to promote both types of mobility. As a means of arriving at a more integrated analysis of social mobility, our first contribution is therefore to account for the dependence between measures of absolute and relative mobility and to explicitly model the parameter that governs the strength of their dependence as a function of covariates within a regression setting. This allows us to provide a more thorough insight into their association, since we are able to analyze how the interrelationship between measures of social mobility changes as a function of economic factors. Our analysis is empirically motivated by a recent large-scale study of Chetty et al. (2014), who calculate and compare measures of absolute and relative mobility in the United States and find that both measures are highly correlated, which leads them to conclude that areas with a high amount of relative social mobility also provide Americans with high rates of absolute mobility.

In an additional initiative towards a more comprehensive analysis of social mobility, and as a second contribution of the present analysis, we add to a recent strand of literature in economics and social science that highlights the importance of analyzing conditional heteroscedasticity in addition to the standard conditional mean analysis. Contrary to the commonly held view that heteroscedasticity is only relevant when it comes to alleviating adverse effects on statistical inference, we follow the works of Downs and Rocke (1979), Western and Bloome (2009), Zheng et al. (2013) and consider its analysis to be an important source of revealing additional information that would otherwise go undetected. Extensions of the conventional regression models, termed heteroscedastic regression models (HRM, Smyth (1989)), variance function regression models (Western and Bloome, 2009), double generalized linear models (DGLM, Smyth et al. (2001); Smyth (2002)) or double hierarchical generalized linear models (DHGLM, Nelder and Lee (1991)), have recently been used in economics and sociology not only to detect violations of standard OLS assumptions, but also for substantive insight. An example of this includes the excess residual variation in income inequality within certain population subgroups that has been interpreted as reflecting unobserved skills or economic insecurity (see, e.g., Western and Bloome (2009)). In general, heteroscedastic regression models are intended to model the residual variation as a function of covariates within a regression setting, instead of treating it as a nuisance only.[3] Although heteroscedastic regression models have led to new empirical insight with respect to the analysis of earnings insecurity and income inequality (Western et al., 2008; Western and Bloome, 2009), research on social mobility has so far neglected the additional information in the data that is revealed by the use of such models. Not only do differences in the rate of social mobility between groups of people or geographic areas matter, but also how access to opportunity and the prospects of upward social mobility vary within groups. In particular, besides knowing the factors that influence the overall level of social mobility, it is also important to look inside different strata of the data and to model the internal variability in order to better understand and analyze the factors that determine why access to social mobility varies within population groups, over time or across space. As with the literature on income inequality, heteroscedasticity with respect to social mobility can be interpreted in several ways: excess residual variation might reflect unobserved skills that are supportive in improving people's economic standing. Moreover, heteroscedasticity, as a measure of variability, can be interpreted as a means to assess the risk or insecurity attached with participating in upward social mobility. For instance, consider an area in which social mobility varies considerably as a function of observable and unobservable factors, so that people do not equally participate in upward social mobility. It is therefore of great interest for academics and policy development alike to analyze why some areas are more alike in terms of providing access to social mobility and how the variability is related to, e.g., gender, race or other personal and regional characteristics. Since the presence of within-group heterogeneity and its analysis can lead to new empirical insight (Zheng et al., 2013), our second contribution aims at investigating conditional heteroscedasticity of social mobility within a regression setting.

From a methodological point of view, and as our third contribution, we add to the empirical literature in economics and social science that is centered on analyzing the entire conditional distribution. In particular, we make use of Bayesian multivariate distributional regression of Klein et al. (2015), which allows us to model not only conditional means and conditional variances, but all distributional parameters of a parametric response distribution within a unified regression setting. Doing so allows for a more realistic prediction of the conditional distribution of absolute and relative social mobility. The model developed by Klein et al. (2015) further extends the univariate modelling of a single response to the multivariate case, where in addition to the modelling of each distributional parameter of the marginals, the dependence between the responses can be modelled as a function of covariates within a regression setting. With respect to the existing approaches, the model presented by Klein et al. (2015) encompasses most of the heteroscedastic regression models proposed in the literature as special cases. Additionally, while most heteroscedastic regression models are restricted in the sense that they are embedded in the framework of Generalized Linear Models (GLM), distributional regression is more flexible since it relaxes the requirement of the response distribution to be a member of the exponential family by allowing for more general distributions. With respect to the modelling of each parameter of a multivariate response distribution, the predictor of the model offers a maximum degree of flexibility since it includes linear terms, non-linear terms such as penalized splines, random effect terms, varying coefficients, and spatial effects. The inclusion of the latter term is especially useful for our analysis as it enables us to explicitly take the spatial dimension of social mobility into account and to graphically assess the ability of the factors that have been discussed in the sociology and economics literature in an effort to capture the spatial patterns of social mobility across the United States. The analysis of local spatial patterns is important as it allows for a better understanding of social mobility and to draw a more precise picture at a regional level, so that measures which aim to improve social mobility can be targeted more precisely (Sharkey and Graham, 2013). Consequently, the model that is used for the present analysis allows us to investigate the questions raised above within a single regression framework. At the time of writing, we are not aware of any study that analyses conditional means, conditional variances, as well as conditional dependence of social mobility within a unified regression framework.

This paper is structured as follows: Section 2 presents the research questions upon which the paper is organized, while Section 3 briefly introduces the regression model. Section 4 gives an overview of the data and Section 5 deals with variable selection and model choice. Estimation results are presented in Section 6. Section 7 concludes.

2. Research questions

In a recent large-scale study, Chetty et al. (2014) investigate social mobility within the United States in great detail. By calculating various measures of both absolute and relative mobility at a highly disaggregated spatial level, Chetty et al. (2014) argue that the question of whether the United States qualifies as the land of opportunity might not be sensible, as the answer depends largely on geographic location. Based on the influential study of Chetty et al. (2014), our analysis sheds light on the following questions:

Q1: What are the strongest predictors of absolute and relative social mobility and how well do these factors capture spatial patterns across the United States?

In order to identify potential factors that may have some explanatory power for social mobility, Chetty et al. (2014) correlate their measures of social mobility with several variables that have been discussed in the sociology and economics literature. In particular, Chetty et al. (2014) find that children who grow up in less segregated areas with lower levels of income inequality, a good primary school system, higher levels of social capital and more stable family structures are more upwardly mobile in comparison to children from areas that lack these characteristics. Although Chetty et al. (2014) stress that these factors cannot be interpreted as causal for social mobility, these variables may well serve as candidate factors that guide future studies in the search for causal relationships.

The correlational analysis of Chetty et al. (2014) provides initial insight into the way that economic variables act on measures of social mobility. However, in order to better understand the mechanisms driving social mobility, it is also important to assess the ability of these variables to capture the spatial patterns across the United States. Consequently, the focus of this research question is on investigating spatial patterns that emerge from heterogeneities that are left unexplained after consideration of the variables discussed in Chetty et al. (2014).

Q2: How much variation of absolute and relative mobility exists within districts and how is this variation related to economic variables?

In addition to substantial spatial variation across commuting zones, Chetty et al. (2014) also report that social mobility greatly varies across counties within commuting zones. In particular, Chetty et al. (2014) find that the variation of intergenerational mobility within commuting zones is almost as high as the variation across commuting zones.

Although Chetty et al. (2014) point out the importance of the variation of social mobility within commuting zones, they do not further investigate the issue. It is also of interest, however, to analyze why some areas are more alike in terms of providing access to social mobility, while the chances in others are distributed more heterogeneously. Following the recent trend of bringing the attention of economists and sociologists to heteroscedastic regression models that allow for economically investigating the conditional variance in addition to the standard conditional mean analysis, we study the within-district variability of social mobility within a regression framework and model heteroscedasticity as a function of economic factors. Doing so allows for studying the distribution of social mobility in more detail and for identifying candidate factors that govern the insecurity attached to participating in upward social mobility.

Q3: Which areas show the highest dependence between absolute and relative mobility, and what are the candidate factors that govern the strength of their interdependence?

As far as the measurement of social mobility and the assessment of the state of opportunity of a society are concerned, measures of absolute and relative mobility have traditionally been analyzed independently from one another, where a definite choice had to be made concerning which form of social mobility to promote and focus on. However, there is an increasing level of awareness in the literature that, aside from the individual analysis of each measure, it is also important to consider each measure's interdependence by assessing whether children can move beyond their parents in absolute income levels and relative economic standing simultaneously (Haskins et al., 2008; Urahn et al., 2012). Although there is no a priori reason to assume that both measures are linked, it may be the case that both measures show some dependence and may even reinforce each other (Reeves and Sawhill, 2014). Indeed, by correlating the two measures, Chetty et al. (2014) report an unweighted coefficient of correlation across commuting zones of -0.68, which leads them to conclude that areas with a high amount of relative social mobility also provide Americans with high rates of absolute mobility.

Although the finding of a high correlation between the two measures of social mobility is an interesting finding in itself, it does not provide insight into the interrelationship between the two measures. Only through an in-depth analysis of their dependence structure, however, can candidate factors be identified that govern the strength of the association between measures of absolute and relative mobility. To further investigate the link between the two measures, we model the coefficient of correlation between absolute and relative mobility within a regression setting. This allows us to provide insight into their interdependence and to analyze how economic factors act on their mutual dependence.

3. Bayesian structured additive multivariate distributional regression

In order to empirically investigate the research questions formulated in Section 2, as well as to integrate all features of the data into one modeling framework, we need a model with the following characteristics: to explore Q1, the model should allow for the explicit modeling of the spatial dimension, as well as for the unobserved heterogeneity in the data. In order to investigate Q2, higher order moments of the response distribution, such as the variance, need to be modeled in a regression setting by allowing them to be set as functions of explanatory variables. Moreover, rather than assuming conditional independence between the two measures of social mobility, the analysis of Q3 requires the modeling of the parameter that governs the strength of their interdependence as a function of covariates within a regression setting.

These considerations suggest using Bayesian structured additive multivariate distributional regression as introduced by Klein et al. (2015), since this model class is flexible enough to integrate all of the requirements into a unified modeling framework. Bayesian structured additive multivariate distributional regression extends the univariate concept of Bayesian structured additive distributional regression recently introduced by Klein et al. (2013), and replaces the one-dimensional response vector with a multivariate one. By assuming a reasonable multivariate response distribution, the model of Klein et al. (2015) allows for the modeling of the parameter that governs the dependence structure between the response variables within a regression setting, along with the other distributional parameters, such as the mean and the variance. This flexibility provides not only insight into the conditional dependence between the response variables, but also allows us to study heteroscedasticity within a regression framework. Additionally, the flexibility aids in avoiding biased and inefficient estimates of covariate effects that may result from neglecting possible dependencies between the variables of interest (Lang et al., 2003). Moreover, rather than an equation by equation analysis, in which estimates of the distributional parameter from the previous regression are used in the current estimation, the joint estimation and analysis of each distributional parameter allows for taking the estimation uncertainty into account which results in more precise estimates and predictions. Comparing the model of Klein et al. (2015) to existing approaches, the univariate modelling of the marginals encompasses most of the heteroscedastic regression models proposed in the literature as special cases, however, with a predictor that offers a maximum degree of flexibility, since the range of modelling alternatives includes linear terms, non-linear terms such as penalized splines, random effects, varying coefficients, as well as spatial effects. With respect to the conditional distribution of the marginal responses, the model of Klein et al. (2015) is less restrictive than models within the Generalized Linear or Generalized Additive Model framework, since the exponential family assumption is relaxed and replaced by more general distributions. As far as multivariate modelling is concerned, the model of Klein et al. (2015) poses an extension of seemingly unrelated regression models (SUR) as introduced by Zellner (1962) for multivariate normal distributions, since it overcomes the problems that result from neglecting the dependence of higher moments or correlations of the response vector on covariates.

To fix ideas, the framework of Bayesian structured additive multivariate distributional regression of Klein et al. (2015) assumes that observations $(\boldsymbol{y}_i, \boldsymbol{x}_i, \boldsymbol{z}_i, s_i)$, $i = 1, \ldots, n$ are given, where $\boldsymbol{y}_i = (y_{i1}, y_{i2})'$ is a 2-dimensional vector of responses, $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ carry the information of categorical and continuous covariates, respectively, and $s_i \in \{1, \ldots, S\}$ denotes the geographical region s that observation i belongs to.[4] Furthermore, let $p\,(y_{i1}, y_{i2}|\boldsymbol{\vartheta}_i)$, $i = 1, \ldots, n$, denote the conditional density of the 2-dimensional response vector $\boldsymbol{y}_i$, where the conditional density is assumed to be dependent on K distributional parameters, i.e., $\boldsymbol{\vartheta}_i = (\vartheta_{i1}, \ldots, \vartheta_{iK})'$. The general idea of Bayesian distributional regression is to model each distributional parameter ϑ_k, $k = 1, \ldots, K$ as a function of covariates within a regressing setting, where each distributional parameter ϑ_k is linked to a semi-parametric structured additive predictor η^{ϑ_k}. In order to ensure that the restrictions on the parameter space of ϑ_k are met, each parameter ϑ_k is linked to the predictor η^{ϑ_k} by a response function $h_k(\cdot)$. As examples, distributional regression allows for modelling the expectations $E(y_{i1}) = \mu_{i1}, E(y_{i2}) = \mu_{i2}$, variances $\text{var}(y_{i1}) = \sigma^2_{i1}, \text{var}(y_{i2}) = \sigma^2_{i2}$, as well as the coefficient of correlation $\text{corr}(y_{i1}, y_{i2}) = \rho_i$ as functions of covariates

$$\eta_i^{\mu_1} = \mu_{i1}, \quad \eta_i^{\mu_2} = \mu_{i2}, \quad \eta_i^{\sigma_1^2} = \log(\sigma^2_{i1}), \quad \eta_i^{\sigma_2^2} = \log(\sigma^2_{i2}), \quad \eta_i^{\rho} = \rho_i/\sqrt{1-\rho_i^2}$$

for the bivariate Normal distribution, with the additional predictor $\eta_i^{n_{df}} = \log(n_{df,i})$ for the degrees of freedom $n_{df} > 0$ for the bivariate Student-t distribution.[5] For each of the K distributional parameters ϑ_k of the 2-dimensional response distribution, the semi-parametric structured additive predictor η^{ϑ_k} has the following structure (Fahrmeir et al., 2013):

$$h_k^{-1}(\vartheta_{ik}) = \eta_i^{\vartheta_k} = \boldsymbol{x}_i'^{\vartheta_k}\boldsymbol{\beta}^{\vartheta_k} + \sum_{j=1}^{p_k} f_j^{\vartheta_k}(z_{ij}^{\vartheta_k}) + f_{geo}^{\vartheta_k}(s_i), \qquad k = 1, \ldots, K. \tag{1}$$

where each of the distributional parameters ϑ_k may possibly depend on a different set of covariates. In Equation (1), $\boldsymbol{x}_i'^{\vartheta_k}\boldsymbol{\beta}^{\vartheta_k} = \beta_0^{\vartheta_k} + \beta_1^{\vartheta_k}x_{i1}^{\vartheta_k} + \ldots + \beta_q^{\vartheta_k}x_{iq}^{\vartheta_k}$ is the parametric part of categorical covariates including the overall intercept $\beta_0^{\vartheta_k}$, $f_j^{\vartheta_k}$ are non-linear smooth effects of the continuous covariates and $f_{geo}^{\vartheta_k}(s_i)$ incorporates the spatial dimension into the model by inducing spatial dependence between observations nested within regions. Relaxing the restrictive assumption of linearity allows for additional insight to be gained into the relationship between covariates and the response, as important features in the data might go undetected if linear regression models are used. Moreover, allowing the covariates to influence the response in a non-linear way is particularly suitable for modeling higher order moments of the response distribution, as the absence of any prior knowledge makes the formulation of a hypothesis difficult with respect to the functional form between, e.g., the variance or the coefficient of correlation and the covariates (Acar et al., 2011).[6] For modeling the unknown functions $f_j^{\vartheta_k}$, Lang and Brezger (2004) and Brezger and Lang (2006) introduced a Bayesian analogue to P(enalized)-splines which were originally proposed from a frequentist point of view by Eilers and Marx (1996). In the frequentist setting, it is assumed that the unknown function $f_j^{\vartheta_k}$ can be approximated by a polynomial spline of degree l_j. The spline is then represented as a linear combination of D_j B-spline basis functions $B_{j,d_j}^{\vartheta_k}$ evaluated at pre-specified knots $z_{j,min} = \zeta_{j,1} < \zeta_{j,2} < \ldots < \zeta_{j,h_j} = z_{j,max}$

$$f_j^{\vartheta_k}(z_{ij}) = \sum_{d_j=1}^{D_j} \gamma_{j,d_j}^{\vartheta_k} B_{j,d_j}^{\vartheta_k}(z_{ij}), \quad i = 1, \ldots, n, \quad j = 1, \ldots, p. \tag{2}$$

where the coefficients $\gamma_{j,d_j}^{\vartheta_k}$ can be interpreted as amplitudes that scale the basis functions $B_{j,d_j}^{\vartheta_k}$ accordingly. To ensure a good fit to the data, Eilers and Marx (1996) suggest using a sufficiently high number of equidistant knots, as well as to simultaneously impose a penalty based on d-th order differences on adjacent B-spline coefficients $\gamma_{j,d_j}^{\vartheta_k}$ in order to prevent $f_j^{\vartheta_k}$ from being too volatile, where d is usually set to d =2. In the Bayesian framework of Lang and Brezger (2004) and Brezger and Lang (2006), priors for the coefficients $\boldsymbol{\gamma}_j^{\vartheta_k} = (\gamma_{j1}^{\vartheta_k}, \ldots, \gamma_{jD_j}^{\vartheta_k})'$ of the non-linear smooth effect $f_j^{\vartheta_k}$ are defined by replacing the difference penalty by first or second order random walks, where the amount of smoothness of $f_j^{\vartheta_k}$ is estimated simultaneously with the regression coefficients $\boldsymbol{\gamma}_j^{\vartheta_k}$.

Similar to Generalized Additive Mixed Models (GAMM) for longitudinal data, where random effect terms allow for relaxing the assumption of independence between repeated observations within a group or cluster, the spatial term $f_{geo}^{\vartheta_k}(s_i)$ in Equation (1) accounts for spatial autocorrelation and acts as a surrogate for unobserved covariates that are not included in the model (Fahrmeir and Kneib, 2011). In order to arrive at a smooth surface, it has to be ensured that the estimated spatial effects $\boldsymbol{f}_{geo}^{\vartheta_k} = \left(f_{geo}^{\vartheta_k}(s_1), \ldots, f_{geo}^{\vartheta_k}(s_n)\right)' = \boldsymbol{Z}_{geo}^{\vartheta_k}\boldsymbol{\gamma}_{geo}^{\vartheta_k}$ of neighboring districts do not differ too much from one another, where $\boldsymbol{\gamma}_{geo}^{\vartheta_k} = \left(\gamma_{geo}^{\vartheta_k}(1), \ldots, \gamma_{geo}^{\vartheta_k}(S)\right)'$ is a vector of regression coefficients that collects all distinct spatial effects for the districts $s = 1, \ldots, S$ and $\boldsymbol{Z}_{geo}^{\vartheta_k}$ is a $(n \times S)$ design matrix that connects an observation i with the corresponding spatial effect, i.e., $\boldsymbol{Z}_{geo}^{\vartheta_k}[i, s] = 1$ if $\boldsymbol{y}_i = (y_{i1}, y_{i2})'$ was observed in district s and 0 otherwise. One way of achieving such spatial smoothness is to assign Gaussian Markov Random Field (GMRF) priors to the coefficients:

$$\gamma_{geo}^{\vartheta_k}(s)|\boldsymbol{\gamma}_{geo}^{\vartheta_k}(-s) \sim N\left(\frac{1}{|N(s)|}\sum_{r\in N(s)}\gamma_{geo}(r), \frac{\tau_{geo}^2}{|N(s)|}\right), \quad s = 1, \ldots, S. \tag{3}$$

where $\boldsymbol{\gamma}_{geo}^{\vartheta_k}(-s)$ is the vector containing all spatial effects except for the one for district s, $|N(s)|$ denotes the total number of neighbors that share a common boundary with district s and the parameter τ_{geo}^2 controls for the variation of $\gamma_{geo}^{\vartheta_k}(s)$ around its expected value. From Equation (3), it follows that the estimated spatial effect $\gamma_{geo}^{\vartheta_k}(s)$ is a weighted average of the neighboring effects, with region s being conditionally independent from all other regions that do not share a common boundary and where the variability of $\gamma_{geo}^{\vartheta_k}(s)$ around its conditional mean inversely depends on the number of neighbors that surround region s. Consequently, assigning Markov random field priors induces a specific correlational structure and ensures spatial smoothness of the regression coefficients $\boldsymbol{\gamma}_{geo}^{\vartheta_k}$, since parameters of neighboring districts are not allowed to deviate too strongly from one another. Therefore, instead of treating the spatial structure in the data as a nuisance that holds little interest on its own, adding $f_{geo}^{\vartheta_k}$ to the predictor in Equation (1) and explicitly modeling and plotting the estimated effects allows us to present insight into the spatial association of social mobility that would otherwise go unnoticed. For a more detailed exposition of structured additive regression models, we refer the interested reader to Fahrmeir and Kneib (2011) and Fahrmeir et al. (2013).[7]

4. Data

For our analysis, we make use of the data provided by Chetty et al. (2014).[8] Using information from federal income tax records of more than 40 million American adult children and their parents, Chetty et al. (2014) compute measures of absolute and relative mobility both at the county and commuting zone (CZ) level.[9] Doing so allows Chetty et al. (2014) to draw conclusions about social mobility in the United States at a highly disaggregated spatial level. In addition to measures of social mobility, Chetty et al. (2014) also provide a number of variables in the form of local area characteristics, mainly at the commuting zone level, that have previously been discussed in the sociology and economics literature.

To construct the different measures of social mobility, Chetty et al. (2014) rank both parents and children within the same birth cohort based on their position in the national income distribution. By examining the relationship between mean child and parent rank across commuting zones, Chetty et al. (2014) find that this relationship is approximately linear, which allows them to model the conditional expectation of a child's rank given his or her parent's rank in a particular commuting zone within the following regression setting:

$$R_{ic} = \alpha_c + P_{ic}\beta_c + \epsilon_{ic} \tag{4}$$

where R_{ic} denotes the national income rank of child i who grew up in commuting zone c. Likewise, P_{ic} denotes the parent's rank in the national income distribution. Chetty et al. (2014) define the slope β_c of the rank-rank relationship in Equation (4) as a measure of relative social mobility that determines how strongly the position of the children in the income distribution depends on the parent's position. Consequently, the lower the β_c, the higher is social mobility in relative terms. Similarly, the intercept α_c is a measure of the expected rank for families with incomes at the lower end of the income distribution. Combining the intercept α_c and slope β_c from Equation (4), Chetty et al. (2014) continue to define absolute social mobility as follows:

$$\bar{r}_{pc} = \alpha_c + p\beta_c \tag{5}$$

where $\bar{r}_{pc}$ is the expected rank of child i who grew up in commuting zone c with parents having a national income rank at percentile p. In their study, Chetty et al. (2014) focus primarily on absolute social mobility for children coming from families that fall below median parent income in the national income distribution and set $p = 25$.[10] Consequently, given the linearity of the rank-rank relationship, $\bar{r}_{25c} = \alpha_c + 25\beta_c$ is a measure of the expected mean income rank of children with parents in the 25-th percentile in the national income distribution.[11] In their study, Chetty et al. (2014) show that the prospects for economic mobility vary considerably across the United States and depend substantially on the location in which children are raised.

As opposed to a full discussion of all variables provided in Chetty et al. (2014), we restrict the number of variables to a subset right from the start. The initial restriction is due to the fact that some of the variables are not available for all commuting zones for which measures of absolute and relative mobility have been calculated. This is especially true for variables in the parent groups that are supposed to reflect the quality of schooling and college education, where, e.g., the variable 'number of colleges per capita' has 214 missing values, whereas the variables 'teacher student ratio' and 'High School dropout rate' have 161 and 482 missing values, respectively. Consequently, in order to ensure a reasonable number of counties and commuting zones for the estimation of the spatial effects, we exclude all variables at the commuting zone level that have 21 or more missing values, as this number turns out to be the ceiling for most variables. Moreover, we focus on the continental United States in our study and exclude Alaska and Hawaii from the analysis. These exclusions leave us with 2,728 counties and 675 commuting zones for which full data is available. Still, aside from the parent group 'College education', at least one variable is available within each of the categories that are investigated more closely by Chetty et al. (2014) in their correlational analysis.[12] Moreover, following Kourtellos et al. (2015), we exclude the variables 'segregation of poverty', 'segregation of affluence', 'Gini coefficient for parent income' and 'fraction middle class' due to multicollinearity. Table 1 shows the set of variables that are used for our analysis.

[Table 1 about here.]

5. Model choice and variable selection

Before presenting the analysis of social mobility, a number of decisions must be made. First, it is necessary to decide on a candidate multivariate distribution that allows for adequately modelling both marginal distributions and for capturing the dependence structure between the two measures of social mobility. The second decision is concerned with variable selection, since there is a large number of candidate factors in the data that possibly influence the two measures of social mobility. While variable selection is already a challenging task in the context of univariate regression analysis, the situation is further complicated by the necessity of selecting a subset of variables for each marginal of the 2-dimensional response distribution, where each of the K distributional parameters themselves may depend on a different set of covariates. Since there is no economic theory that dictates the selection of covariates across the entire response distribution, selecting an appropriate set of variables for each distributional parameter is crucial with respect to reliable and realistic predictions of social mobility.

With respect to the scale of the data, we follow Chetty et al. (2014) and initially standardize both measures of absolute and relative mobility, as well as all covariates in the data to have a mean of 0 and a standard deviation of 1. This standardization allows for making the results of this analysis comparable to those presented by Chetty et al. (2014). As candidate response distributions for measures of absolute and relative mobility, we consider the bivariate Normal and the bivariate Student-t. Concerning the choice of an appropriate subset of covariates for each distributional parameter of the bivariate Normal and the bivariate Student-t distribution, i.e., for

$\eta^{\mu_{abs}}$, $\eta^{\mu_{rel}}$, $\eta^{\sigma^2_{abs}}$, $\eta^{\sigma^2_{rel}}$, η^{ρ} and $\eta^{n_{df}}$, the fact that no underlying economic theory exists which provides guidance in selecting a reasonable set of economic variables for higher order moments of the response distribution, leads to the use of stepwise forward variable selection. Using this procedure, the variable that improves the model the most in terms of the deviance information criterion (DIC) (see Spiegelhalter et al. (2002) for details) is iteratively included, where the procedure is stopped if none of the remaining variables leads to an improvement of the model anymore. For both bivariate distributions, the most influential covariates are selected, first for the means and then proceeding on to variances and to the coefficient of correlation, ending with the degrees of freedom for the bivariate Student-t. In order to assess and compare the fit to the data under different distributional assumptions, we follow Klein et al. (2013) and use the concept of quantile residuals of the marginal distributions (see Dunn and Smyth (1996) for details), as well as the DIC. Quantile residuals allow for the assessment of the distributional assumption made during the modeling process and for differentiating between competing distributions by graphically comparing the fit of the residuals with the bisection line; the closer the residuals agree with the line, the better is the fit to the data. The virtue of quantile residuals, besides being a convenient graphical tool to assess the distributional assumption, is that they are robust with respect to the exact specification of the predictor and, consequently, the selected covariates (Klein et al., 2013). The results for the bivariate Normal and the bivariate Student-t distribution are shown in Figure 1, where quantile residuals are computed for each marginal distribution.[13]

[Figure 1 about here.]

While both distributions provide a reasonable fit to the data for more central parts, residuals outside the range of -3 to 3 clearly deviate from the line for the Normal distribution. In contrast, due to its heavier tails, the Student-t distribution is better able to capture the overall shape of the residuals, especially for more extreme observations, even though the residuals of the Student-t marginals deviate slightly from the bisection line in the lower part of the distribution for absolute mobility. The ability of the bivariate Student-t distribution to better describe the data is also reflected in the values of the DIC, as shown in Table 2.

[Table 2 about here.]

Even though each marginal distribution can each be approximated by a Student-t distribution, using a bivariate Student-t distribution for the analysis implies that the dependency between measures of absolute and relative social mobility can reasonably be captured by the coefficient of correlation. However, this assumption might be too restrictive in the sense that the dependency between the two measures is governed by a parameter other than the coefficient of correlation of the bivariate Student-t distribution. Consequently, it is important to separate the univariate marginal distributions from the dependence structure, so that the mutual dependence between measures of absolute and relative social mobility can be investigated. The framework of copulas provides a way of doing so, in that it allows to rewrite any joint multivariate distribution function in terms of univariate marginal distributions and a function that governs the dependence structure between the variables, which is referred to as the copula (see Joe (2014) or Nelsen (2007) for an introduction to copulas). The initial step is to use the conditional cumulative density function of the univariate Normal and the Student-t distribution, where each distributional parameter is modeled as a function of covariates, and to transform both measures of absolute and relative mobility to the unit interval, so that each marginal is distributed on $[0, 1]$. Doing so allows for fitting several copulas to the data in a second step and to choose the one that best describes the dependency between absolute and relative social mobility. Figure 2 shows the scatterplots and the histograms of the transformed data.

[Figure 2 about here.]

Similar to the results based on quantile residuals and the DIC, Figure 2 shows that the Student-t is a reasonable approximation to both measures of mobility, as the transformed marginals are approximately uniformly distributed, as opposed to the transformed marginals of the Normal distribution which show some deviation. Also note that the two measures exhibit dependence in both tails, as well as towards the (0,0) and (1,1) corners, which is an indication that the Student-t copula may be a reasonable choice for modeling the dependence structure. In order to investigate whether the coefficient of correlation of the bivariate Student-t distribution adequately describes the dependency between the two measures of social mobility, we rely on the value of the Akaike information criterion (AIC) (see Akaike (1974) for details) as a means to distinguish between competing copulas. As candidates within the class of elliptical copulas, the Normal and the Student-t copula are considered. Within the Archimedean family, we restrict the analysis to the Frank and the Clayton copula.[14] Table 3 displays the results of the AIC for the different choices of the copulas.[15] To compare the fit of the copulas with respect to the different specifications of the marginal distributions, the results for the Normal distribution are also included.

[Table 3 about here.]

Comparing the different specifications of marginal distributions and copulas, Table 3 shows that assuming a Student-t distribution for both marginals, as well as a Student-t copula to capture the dependence between the two measures of social mobility provides the best fit. Consequently, the bivariate Student-t provides a reasonable fit to the data, not only in terms of marginal distributions, but also with respect to the modeling of the dependence structure between measures of absolute and relative mobility. Therefore, the bivariate Student-t distribution is used for the analysis, with Table 4 showing the specification of the additive predictors and the selected covariates for each distributional parameter based on stepwise forward variable selection using the DIC as described above.

[Table 4 about here.]

6. Analysis of intergenerational mobility

Similar to the opinion expressed in Chetty et al. (2014), it is necessary to stress that the results presented in the following are to be understood in a more descriptive and exploratory sense, rather than to give specific policy recommendations. Since we use data on the level of counties and commuting zones, we would also like to make the reader aware of an issue that is related to the analysis with respect to aggregated data. According to the ecological fallacy, the researcher should be aware that there is a logical sophism inherent in making causal inference from group data to individual behavior (Robinson, 1950; Selvin, 1965).[16] That is, using group data from a particular area to infer individual behavior in that area might lead to erroneous conclusions, since the relationship found in aggregate data might not hold for individuals. Moreover, it must be stressed that although a strong statistical relationship exists, it does not necessarily imply a causal relationship between the economic variables and measures of intergenerational social mobility (Chetty et al., 2014). It is possible that other factors are at play which are responsible for the seeming relationship between any of the economic variables and social mobility, as the dependence may result from unobserved variables that induce a certain correlation which may or may not hold after taking additional variables into account (Badel and Maues, 2014).[17] In contrast to Chetty et al. (2014), however, including the spatial effects term allows us to control for numerous unobserved differences across districts, since $f_{geo}^{\vartheta_k}(s)$ acts as a surrogate for unobserved covariates that are not included in the model. Moreover, explicitly modelling and plotting the spatial effects term enables us to investigate spatial patterns that emerge from heterogeneities that are left unexplained after taking covariates into account. In order to investigate the ability of the covariates to explain the heterogeneity of absolute and relative mobility across the United States, all moments of the bivariate Student-t distribution are estimated in a reduced model, where only the spatial effect effect term $f_{geo}^{\vartheta_k}(s)$ is included in $\eta^{\mu_{abs}}$, $\eta^{\mu_{rel}}$, $\eta^{\sigma^2_{abs}}$, $\eta^{\sigma^2_{rel}}$, η^{ρ}. The estimated effects are then compared to those of the full model, for which the most informative covariates have been included in the semi-parametric predictor η^{ϑ_k}.[18] Significance maps further enhance the detection of spatial patterns by classifying the estimated spatial effect into one of three categories; the spatial effect is classified as insignificant at the 95% level and the corresponding district is colored in grey if the credible interval includes zero. Districts with significant positive effects are colored in white, whereas districts with significantly negative effects are colored in black. By comparing and plotting the effects of $f_{geo}^{\vartheta_k}(s)$ for the reduced and the full model, we are able to assess the explanatory power of the covariates that are included in the full model and to graphically investigate the remaining spatial patterns that assist in identifying additional factors that have to be included in the model in order to capture the unexplained heterogeneity in the data.[19]

6.1. Modeling conditional means of social mobility

Similar to Chetty et al. (2014), the main focus of this section lies in analyzing the factors that influence measures of absolute and relative social mobility in a conditional mean regression setting. As far as the discussion of variable selection and the estimated effects of covariates are concerned, we restrict our presentation to only a few selected effects.[20] In particular, we relate our results to the findings of Chetty et al. (2014), as well as compare them with those presented in Kourtellos et al. (2015), who use Bayesian model averaging in order to identify a set of robust predictors for social mobility among the variables provided by Chetty et al. (2014).

What appears from the number of selected covariates presented in Table 4, both for absolute and relative mobility, is that social mobility is governed by a complex data generating process, since the number of selected covariates exceeds the number of influential factors that have been identified by Chetty et al. (2014). This is in accordance with the results presented in Kourtellos et al. (2015), who report a posterior mean of a model size of 14 in order to capture the most prominent aspects of absolute mobility. Comparing our results for absolute mobility with those reported in Chetty et al. (2014), shows that our variable selection procedure generally identifies the same parent groups of variables that Chetty et al. (2014) identify as having the strongest influence on social mobility: 'segregation', 'income inequality', 'school quality', 'social capital' and 'family structure'. Moreover, Table 4 shows that all of the variables that have been identified by Chetty et al. (2014) as exhibiting the highest influence on absolute and relative mobility have also been selected for our models (apart from the High School dropout rate for absolute mobility since this variable was excluded from the analysis due to a large number of missing values and the Gini for the bottom 99% of the income distribution for relative mobility). Comparing our results to Kourtellos et al. (2015) shows that the groups and the selected variables therein also generally agree with the those identified by Bayesian model averaging.[21] Similar to Kourtellos et al. (2015), our results indicate that local labor market conditions, such as the fraction working in manufacturing, also exhibit an influence on absolute mobility as nearly all of the variables within this group have been selected. Moreover, tax-related issues, such as state income tax progressivity, or the quality of schooling, represented by test score percentiles, also have an influence on absolute mobility.[22]

With regards to variable selection for relative mobility, Table 4 shows that, similar to Kourtellos et al. (2015), the two measures of social mobility are influenced by a slightly different set of covariates. Although the parent groups of variables are similar for both measures, Table 4 shows that within these groups, the selected variables differ. For example, variables in the parent group 'migration' seem to have a greater influence on relative mobility, since both the migration outflow rate, as well as the fraction of foreign born residents have been selected, compared to only the migration inflow rate for absolute mobility. As for absolute mobility, quality of schooling seems to exhibit a strong influence on relative social mobility, as both the test score percentile, as well as school expenditures per student have been selected as important covariates. Furthermore, tax related issues matter more for relative mobility than for absolute, as indicated by the variables denoted state EITC exposure and state income tax progressivity. In opposition to Chetty et al. (2014), but in line with Kourtellos et al. (2015), our results indicate that it is not only the Gini for the bottom 99% of the income distribution that is relevant for social mobility, but also the income share of the top 1%, since this covariate has been selected for both measures of social mobility.

[Figure 3 about here.]

[Figure 4 about here.]

Comparing the estimated effects with the existing literature, which has traditionally assumed a linear relationship, Figures 3 and 4 show that measures of mobility depend in an explicit non-linear way on some of the covariates.[23] For absolute mobility, this is especially true for the fraction with short commuting times, for the fraction religious, the household income per capita, the fraction of adults married, the teenage labor force participation rate, as well as income segregation. While the effect of income segregation has generally been found to be negatively associated with absolute mobility,

Figure 3 shows a pronounced U-shaped effect where the effect first decreases absolute mobility, before it increases again. For relative mobility, the labor force participation rate, the fraction of African American residents, the fraction of adults divorced and the fraction with short commuting times have a strong non-linear influence. Also note that the way in which the fraction of children with single mothers acts on the response is different for absolute and relative mobility; while the effect is almost linear for absolute mobility with some flattening for very high percentages, the effect on relative mobility is clearly non-linear. While this first causes relative mobility to decrease, there seems to be some threshold after which the effect levels out. The effect after this fraction is accompanied with an increasing uncertainty attached to the estimated effects so that reliable statements beyond this point cannot be made.

In order to assess the ability of the selected covariates to capture the spatial patterns of absolute and relative mobility across the United States, we plot the significance maps of the estimated spatial effects of the reduced and the full model in Figure 5. Investigating the panels of the reduced models in Figure 5 shows that the models reasonably pick up the spatial patterns of absolute and relative mobility. Accordingly, the American Dream is still alive in terms of absolute mobility in the Great Plains, as well as on the West Coast and in some parts of the Northeast. In contrast, rates of mobility for people growing up in the Southeast and in the Rust Belt are particularly low for both measures. Growing up in rural areas generally provides children with higher rates of absolute mobility compared to children growing up in urban areas.

[Figure 5 about here.]

Comparing the reduced with the full models shows that the selected covariates leave heterogeneities for both absolute and relative mobility largely unexplained. While the covariates have some explanatory power for relative mobility in the Great Plains region of the United States, the bottom right panel of Figure 5 shows that relative mobility remains unexplained with the available covariates, especially for those parts in which rates of relative mobility are particularly low, such as in the Eastern and Southeastern regions of the United States. The bottom right panel of Figure 5 also shows that covariates are not well suited to pick up the spatial patterns of high rates of relative mobility, as spatial differences in social mobility in large parts of mobile areas, such as on the West Coast, remain unexplained.

With regards to the full model of absolute mobility, Figure 5 shows that rates of absolute mobility are considerably higher than what can be explained with covariates in highly mobile areas, such as the Great Plains. Although the full model captures some of the heterogeneity of absolute mobility in parts of California, Nevada, Arizona and New Mexico, districts in which rates of mobility are particularly low, such as in the Southeastern region of the United States and around the Rust Belt, exhibit rates of social mobility that remain largely unexplained. Consequently, the panels of the full models in Figure 5 show that the available covariates are not able to account for regional differences in social mobility and that additional economic factors need to be identified.

6.2. Modeling within-district variability of social mobility

While empirical models in economic and sociological research have generally been developed within a conditional mean regression setting, recent advances in statistics have led to an expansion of interest also into studying covariate effects on heteroscedasticity. Whereas conventional regression models have treated heteroscedasticity as a nuisance that has to be adjusted for in order to alleviate its adverse effects on statistical inference, extensions of these models view violations of homoscedasticity as potentially important sources of additional information in the data that is systematically related to covariates within a regression setting (Zheng et al., 2013). In particular, while economic and sociological theory has traditionally been based on hypotheses concerning between-group differences, heteroscedastic regression models extend the analysis to analyze differences within-groups as well (Western and Bloome, 2009), where groups can be thought of as representing different categories of covariates (e.g., gender, race or geographic areas in spatial statistics and econometrics). As an illustration, consider gender as a categorical covariate that has two groups, male and female. In a standard conditional mean setting, regression coefficients describe differences in group means, e.g., expected differences in monthly salaries between men and women.[24] In addition to analyzing these between-group differences, heteroscedastic regression models extend the analysis to within-group differences, i.e., testing heterogeneity within groups, for example within men and within woman, for systematic differences. In other words, covariate effects in conditional mean regression account for deviations of the group sample means from the overall mean of the response (between-group differences), while covariate effects in heteroscedastic regression models also explain how the variability of the response around group means changes as a function of covariates within groups (within-group difference) (Zheng et al., 2011). Consequently, parallel to studying between-group differences within a conditional mean regression setting, the analysis of within-group heterogeneity modelled as conditional heteroscedasticity yields a more complete picture of the response variable (Zheng et al., 2013). With respect to giving an economic interpretation of within-group differences in the form of heteroscedasticity, the literature on income inequality has offered the interpretation of heteroscedasticity as reflecting the influence of unobserved or hidden heterogeneity in the form of luck (Jencks et al., 1972), skill, such as intrinsic ability, work effort and school quality (Juhn et al., 1993; Lemieux, 2006), or as measuring income risk and insecurity (Western et al., 2008; Western and Bloome, 2009).

Being among the first to apply heteroscedastic regression to the analysis of social mobility, this section analyzes the within-district variability of absolute and relative social mobility. In their study, Chetty et al. (2014) find that the variation of intergenerational mobility within commuting zones is almost as high as the variation across commuting zones which makes it particularly suitable for our analysis, since the analysis of the conditional variance is of special importance in cases where within-group variability is similar in magnitude to between-group differences (Western and Bloome, 2009). With respect to its interpretation, we regard heteroscedasticity as reflecting the insecurity attached to participating in upward social mobility or, similarly, as the risk of being excluded from sharing the proceeds of upward social mobility. The analysis of within-district variability is of interest because the decision of people to remain in or to move to a specific district may not be driven only by the overall level of social mobility, but also by the chances to participate in social mobility. For instance, people might prefer living in areas where the fluctuation around the mean level of social mobility is comparatively low, so that the risk of being excluded from the opportunity of climbing the income ladder is low. With respect to modelling covariate effects, the model used in our analysis is more flexible than existing approaches that model conditional heteroscedasticity, as it allows for including linear terms, non-linear terms, as well as spatial effects. In contrast to existing studies, the inclusion of the latter term is especially interesting, as the explicit modelling and plotting of the spatial effect allows us to graphically identify those areas in which the risk attached to social mobility is particularly high, as well as to assess the explanatory power of the covariates that are included in the model to capture within-district variability of social mobility.

With respect to the estimated covariate effects that influence the within-district variability of absolute social mobility, Figure 6 shows that racial segregation, as well as the local tax rate increase within-district variability of absolute mobility, while all other covariates tend to reduce the variation within districts.

[Figure 6 about here.]

In line with the literature that has identified family as being a crucial element for the development of children (see e.g., Becker (1991); Murray (1984, 2012), panel (1) of Figure 6 shows that more stable family structures, as indicated by the fraction of adults who are married, leads to a more homogeneous access to social mobility.

The fact that the Gini for the bottom 99% has been selected as an influential factor suggests, that in addition to the mean regression case, there is a Great Gatsby Curve for the variance as well.[25] The Gini for the middle class exhibits a non-linear influence on the within-variability of absolute mobility, where a higher inequality for the bottom 99% with respect to income first increases the variance, before balancing out access to absolute social mobility. This presents an interesting extension of previous studies that have investigated the Great Gatsby Curve for the mean only.

Concerning its influence on the mean level of absolute mobility, Chetty et al. (2014) find that the local tax rate, which consists predominantly of property taxes used for financing schools, is positively associated with absolute mobility. However, despite this positive influence on absolute social mobility for the conditional mean, panel (3) of Figure 6 shows that the local tax rate leads to an increase in the variability, so that the distribution, and hence the risk in participating in social mobility, is more spread out. One possible explanation for the effect of the local tax rate on increasing variability of social mobility might be related to its role for local governments funding public education. Since public schools in the U.S. are primarily financed through local property taxes, the best public schools tend to be in high-income areas; hence children with wealthier parents tend to go to the best schools. Therefore, the higher the tax rate, the wealthier an area and the more money is spent on education in that area, so that access to high quality education, and hence access to social mobility, is limited for children in low income areas. With respect to the effect of the fraction of workers that is employed in the manufacturing industry, panel (4) shows that it leads to a more equal access to absolute mobility, with the effect leveling out after a certain threshold. This decreasing effect is in line with findings in the literature for the mean regression case (e.g., Wilson (1996)) which suggest that the availability of jobs in certain areas of the labour market, such as the manufacturing sector, can help lower-skilled workers to move up in the income distribution (Chetty et al., 2014). Similar to the mean regression case, racial segregation has a strong influence, as indicated by the magnitude and significance of the estimated effect. As shown in panel (5) of Figure 6, the more racially segregated an areas is, the more unequal is access to social mobility and the higher is the risk of not participating in upward social mobility. As noted by Chetty et al. (2014), racial segregation may hinder access to social mobility by reducing the exposure of disadvantaged individuals to successful peers and role models, as well as by reducing intercultural exchange with people from different nationalities.

Investigating the effects of the covariates on the within-variability of relative mobility, Figure 7 shows that the parent group 'segregation' has a strong influence, since the fraction of African American residents, as well as racial segregation have been identified as important covariates. Moreover, the distribution of income represented by the portion of income that accrues to the top 1 percent is also an important economic factor for within-district variability of relative mobility.

[Figure 7 about here.]

Concerning the influence of the percentage of African Americans, Chetty et al. (2014) note that social mobility is lower in areas with larger African-American populations. A similar effect seems to be prevalently attached to the risk in participating in relative mobility. However, Chetty et al. (2014) note that the effect of the percentage of African Americans is to be understood in a place-level race effect instead of an individual-level race effect: rather than African American children having lower incomes than Caucasian children, areas with large African American populations have lower rates of upward social mobility for children of all ethnicities. Racial segregation, as with absolute mobility, increases the risk of social mobility within an area. However, the effect of racial segregation shows a behavior consistent with a U-shaped curve compared to that of absolute mobility. In contrast to the variability of absolute mobility, where the distribution of income within the middleclass is more important, inequality between the top 1% and the remainder of the income distribution matters more for the risk attached to relative social mobility. As for absolute mobility, there seems to be a Great Gatsby Curve for the variance of relative mobility: as inequality between the top 1% and the remainder of the population increases, relative mobility first becomes more homogeneously distributed before the risk of participating in the overall level of relative mobility increases. However, due to the uncertainty attached to the estimated effect, as indicated by the wide credible interval, the estimated effect for the portion of income that accrues to the top 1% needs to be interpreted carefully.

Inspection of the spatial patterns of the within-district variability of the reduced models presented in the left panels of Figure 8 shows that social mobility varies considerably within districts. In particular, spatial variability of absolute mobility is high not only in large parts of the Great Plains, but also in areas of the Southeast. In contrast, rates of absolute mobility are distributed more homogeneously on the West Coast and in the Northeastern regions of the United States. With respect to the spatial patterns of relative mobility, Figure 8 shows that the access to mobility is distributed more heterogeneously on the West Coast, while rates of relative mobility are less variable around the Rust Belt.

[Figure 8 about here.]

Comparing the full models in the right panels of Figure 8 to the mean models shows that the selected covariates are better able to capture the spatial patterns of within-district variability, both for highly varying districts, as well as for districts where access to social mobility is distributed more homogeneously. However, there are still some highly varying districts where spatial heterogeneity remains unexplained. While the economic variables are able to account for most of the spatial variability of absolute mobility in the Western, Eastern and Southeastern regions of the United States, within-district variability remains considerably high in parts of the Great Plains, as well as in parts of Texas. For within-district variability of relative mobility, Figure 8 shows that the selected covariates explain most of the variability in highly variable districts, such as on the West Coast. Interestingly, both full models for absolute and relative mobility in the right panels of Figure 8 show a similar pattern of unexplained heterogeneity in parts of the Great Plains, as well as in parts of Texas. As far as the ability of the selected covariates is concerned for explaining the spatial patterns in those areas where absolute and relative mobility are distributed more homogeneously (black districts), the right panels of Figure 8 show that regions in the Northeastern of the United States still show some heterogeneity for absolute mobility that is considerably lower than what can be explained with the selected covariates.

6.3. Modeling the dependence between measures of social mobility

Traditionally, studies of social mobility have focused on either absolute or relative mobility. While absolute mobility captures the idea of growing prosperity for all, where everyone participates in the proceeds of economic growth, relative mobility is concerned with equality of opportunity, with a fluid and meritocratic society providing everyone with equal chances of climbing the income ladder, independent of his or her family background. Typically, societies show a combination of the two, where it is generally possible that high rates of absolute mobility are accompanied by low rates of relative mobility. However, a society in which children are financially better off than their parents, but in which high rates of income inequality make it difficult for children to climb the income ladder, as suggested by the Great Gatsby Curve, is an unfair society, compared to an ideal one which provides growth and prosperity, as well as fluidity and fairness (Reeves and Sawhill, 2014). Although there is no a priori reason to assume both measures to be linked, it is also possible that both measures reinforce each other: higher rates of absolute mobility, resulting from overall economic growth, may lower the barriers to participate in relative mobility, and hence, making it easier to provide equality of opportunity (Reeves and Sawhill, 2014). In any case, understanding the interdependencies between the two, i.e., how high rates of absolute mobility may be reconciled with low relative rates, or why children are able to move beyond their parents in both absolute and relative terms, is essential for policy makers to better balance the need to promote both types of mobility: if measures that aim at improving social mobility in one dimension are taken, but the interdependency between the two is neglected, policy makers could end up with having reduced absolute social immobility at the cost of higher relative social immobility, or vice versa. Based on these considerations and as a step towards a more integrated analysis of social mobility, this section analyzes the pathways through which absolute and relative mobility are linked by relating the coefficient of correlation between the two measures to economic variables within a regression setting.[26]

To empirically motivate our analysis, Chetty et al. (2014) report an unweighted coefficient of correlation of -0.68 between the two measures, which leads them to conclude that areas with high rates of relative social mobility tend to provide children from low income families at $p = 25$ with better outcomes in terms of absolute income. Chetty et al. (2014) further investigate the correlation between absolute and relative mobility across the parent income distribution and find that the coefficient of correlation reaches 0.2 for the wealthiest families and nearly -0.8 for the poorest. Moreover, both measures are uncorrelated for families at the 85-*th* percentile of the parent income distribution and for the top 15%, higher relative mobility actually leads to a decline in absolute mobility (Chetty et al., 2014).

Before we turn to the analysis of the economic factors that govern the strength of the dependence between measures of absolute and relative mobility, we begin with a descriptive spatial analysis of the coefficient of correlation for the reduced model. This preliminary analysis provides a description of those areas where both measures show the strongest dependence, with Figure 9 presenting the results.

[Figure 9 about here.]

The geographic pattern of the coefficient of correlation of the reduced models show that the strength of the dependence between measures of social mobility varies considerably across the United States. In particular, Figure 9 identifies regions around the Rust Belt, as well as parts of the Eastern and Southeastern regions of the United States to have a strong negative effect on the coefficient of correlation, whereas districts in Florida and on the West Coast show a positive effect. Note, that the dark areas in the Eastern and Southeastern regions, as well as those around the Rust Belt generally provide Americans with low rates of social mobility in both dimensions. However, dark areas, such as in parts of the Great Plains, may also indicate that children are able to surpass their parents in absolute and relative terms, as the coefficient of correlation is not able to distinguish between the two scenarios of either high or low rates of social mobility for both measures. Consequently, careful inspection of Figure 9 in comparison with the reduced models in Figure 5 is crucial for the interpretation of the results. Also note that measures of social mobility in light-colored areas generally act in opposite directions, meaning that although rates of social mobility in one dimension are low, these areas still provide Americans with high rates of mobility in the other dimension. This geographic pattern is found mainly in Arizona, New Mexico, California, Oregon, Washington, Florida, in parts of Texas, in parts of South Dakota, as well as in parts of Minnesota and Iowa.

We now turn to the estimated effects of the covariates that act on the coefficient of correlation, with Figure 10 providing the results. [27] For the purpose of comparison, we have also included the unweighted coefficient of correlation of -0.68 between the two measures as reported by Chetty et al. (2014) which is indicated by the dashed line. Note, that while Chetty et al. (2014) have assumed the dependence structure between absolute and relative mobility to be constant for each observation, we provide a more detailed analysis, as the coefficient of correlation is permitted to change as a function of the covariates.

[Figure 10 about here.]

Figure 10 shows that variables within the parent groups that reflect 'segregation' (racial segregation), 'income distribution' (top1percent, household income per capita), 'family structure' (fraction of adults married), 'migration' (fraction of foreign born), 'local labor market characteristics' (labor force participation rate, fraction working in manufacturing) and 'social capital' (fraction religious) have been identified as having an influence on the dependence structure, whereas none of the variables within the parent groups 'tax' or 'K-12 education' have been selected. Comparing the selected variables in Figure 10, as well as the parent groups of variables to those that Chetty et al. (2014) have found to exhibit the strongest influence on the mean level of both measures shows that the parent groups that reflect 'segregation', 'social capital', 'income distribution' and 'family structure' intersect. Comparing the estimated effects to the unweighted coefficient of correlation of -0.68 of Chetty et al. (2014), shows that the coefficient of correlation is not constant across the domain of the selected covariates, but changes as a function of the economic variables. Moreover, some of the covariates influence the coefficient of correlation in a clearly non-linear way, such as the fraction of foreign born, the fraction of adults married or top 1 percent, so that assuming a linear functional form of the estimated effects would be a misspecification with respect to the way in which economic variables act on the mutual dependence between absolute and relative mobility.

According to Figure 10, the fraction of foreign born reduces the negative correlation between absolute and relative mobility and consequently leads to a weakening of their dependence. Indeed, with an increasing fraction of foreign born, both measures become increasingly uncorrelated with the coefficient of correlation gradually approaching zero. Interestingly, while initially being negatively correlated, the nature of the dependence between the two measures of social mobility changes as the share of income that is accrued by the top 1 percent increases, with a coefficient of correlation of about -0.5 changing to nearly perfect positive dependence. Since income inequality in the United States over the last three decades was primarily driven by an increase in the top 1 percent (see Piketty and Saez (2003)), this measure is of particular interest to policy makers.

Similar to panel (2), panel (3) of Figure 10 shows that the fraction of married adults leads to a reverse in the direction of correlation, as the initial positive dependence between the two measures changes into a negative one. As far as the direction and the influence of the other selected covariates on the coefficient of correlation is concerned, Figure 10 shows that while participation in the labor force indicates that the negative correlation slightly decreases, the fraction religious, racial segregation, household income per capita, as well as the share employed in manufacturing further increase the negative correlation between measures of absolute and relative mobility, while the non-linearity of the share employed in manufacturing indicates that it decreases the negative dependence structure for very high shares.

[Figure 11 about here.]

With regards to the spatial effects of the full model, Figure 11 shows that covariates are able to account for the dependence between the two measures in areas that show high rates of relative social mobility and low rates of absolute mobility, e.g., in Florida and in the more Western and Central regions of the United States. The heterogeneity in districts around the Rust Belt, where low rates of mobility in both dimensions are prevalent, is also modeled reasonably well by the covariates. However, some of the spatial heterogeneity remains unexplained in the Southeastern and Southern regions for instance, as well as in parts of Texas.

7. Concluding remarks

Social mobility and the lack thereof is a multilayer problem that is of high social relevance and has long been a topic that has equally attracted the interest of academics and policy makers (Solon, 1992).[28] Social mobility is relevant because it implies equality of opportunity, where everyone is given equal chances of improving his or her socio-economic position through individual effort, hard work and talent, irrespective of his or her birth circumstances. This is not only important for economic prosperity, but is also an attractive political issue, since it is likely to affect social justice: a society in which people have the perception that they can personally pull the strings that lead to an improvement of their economic and social situation, clearly has the potential of fostering social cohesion, as compared to a society with a rigid social hierarchy, in which individual opportunities are decided at birth (Aldridge, 2001). It is therefore of great importance to policy makers and local authorities to understand the mechanisms driving social mobility if corrective actions are to be taken.

This paper provides an extension of the influential study by Chetty et al. (2014) who present an in-depth look at social mobility within the United States by illustrating that social mobility is not equally distributed across the United States, but varies considerably by geographic location. Investigating social mobility not only within a conditional mean setting, but by extending the analysis to conditional variances, as well as to the conditional dependence between measures of absolute and relative mobility, we are able to draw a more complete picture of the state of economic mobility within the United States. Moreover, and in contrast to the existing literature, we relax the assumption of linearity of the estimated effects and show that measures of social mobility depend in an explicit non-linear way on some of the covariates across the conditional distribution. By investigating the informational content of the economic variables provided by Chetty et al. (2014) to capture the spatial patterns of social mobility across the United States, our results for the conditional mean case suggest that even though these factors contribute to a better understanding of its underlying mechanisms, they leave spatial heterogeneities of social mobility largely unexplained.

Following the recent trend of bringing the attention of economists and sociologists to heteroscedastic regression models that allow for economically analyzing heteroscedasticity in addition to the standard conditional mean analysis, this study analyses the within-district variability of social mobility within a regression framework by modelling heteroscedasticity as a function of economic factors. Our analysis shows that the variability of absolute mobility is high in the Great Plains, but also in parts of the Southeast, whereas the variability of relative mobility is high mainly on the West Coast. In contrast to conditional mean models, the selected covariates are better suited to capture the variability within districts. However, spatial heterogeneities remain unexplained in some areas where rates of mobility vary considerably. Moreover, our results suggest that in addition to the case of the mean regression, there is a Great Gatsby Curve for the variance as well so that higher rates of income inequality do not only reduce the overall level of social mobility, but also increase the risk attached to participate in social mobility.

In the final section of our analysis, we relax the assumption of independence between measures of absolute and relative mobility, and contribute to the growing literature that perceives the phenomenon of social mobility as a multivariate one. By explicitly taking the dependency between both measures into account and by modeling the parameter that governs the strength of their interdependence as a function of covariates within a regression setting, we provide insight into the way in which both measures interact and show how the coefficient of correlation between absolute and relative mobility is influenced by economic variables. The spatial pattern of the coefficient of correlation indicates that measures of social mobility show the highest correlation in the Eastern and Southeastern regions of the United States, as well as around the Rust Belt. In particular, the strength of the coefficient of correlation between the two measures is strongest in those areas in which the highest and lowest levels of both absolute and relative mobility occur. Moreover, our analysis shows that the coefficient of correlation is non-constant across the domain of the selected covariates and varies considerably according to the values of economic variables. Even though our analysis is only a preliminary step towards a more integrated analysis of social mobility, we nevertheless hope that we can stimulate further research in this direction in an effort to arrive at a more complete picture of social mobility.

There are several possibilities for extending the current analysis for future research. Although much of the scholarly work has focused on the analysis of upward social mobility, less effort has been devoted to downward social mobility. As far as equality of opportunity and the American Dream are concerned, it is clear that the factors which influence upward social mobility are of primary interest to policy makers. However, in order to better understand the mechanisms driving social mobility, a more thorough analysis of downward social mobility and the factors that determine why later generations are falling behind their parents in both absolute income and relative rank is equally important. Consequently, extending the analysis of social mobility by investigating downward social mobility and its driving factors poses an interesting and promising extension for future research. Another way to develop our analysis in future research is to further investigate the spatial patterns of social mobility, as our results show that the available economic factors are not able to fully account for regional differences in social mobility. Based on the analysis of the spatial effects, our results can guide future studies in identifying additional economic factors that capture the remaining spatial patterns, especially in those areas where rates of social mobility are particularly low. An additional opportunity for future research comes from allowing the economic variables analyzed in Chetty et al. (2014) to vary across space. When modeling the economic variables, we have implicitly assumed that the way in which these factors act on measures of absolute and relative mobility is homogeneous across all districts. However, the effect of one or more variables may vary from district to district. Geographically weighted regression models allow regression coefficients to differ regionally from their global values and assume that the spatial heterogeneity is explained by the space varying regression coefficients (Fotheringham et al., 2002). Since social mobility is a local phenomenon, the analysis of spatially varying covariates is especially interesting for policy makers when it comes to the design and implementation of policy measures, as this analysis allows for a more regional view on social mobility.

References

Acar, E. F., Craiu, R. V. and Yao, F.: 2011, Dependence calibration in conditional copulas: A nonparametric approach, *Biometrics* **67**(2), 445–453.

Akaike, H.: 1974, A new look at the statistical model identification, *IEEE Transactions on Automatic Control* **19**(6), 716–723.

Aldridge, S.: 2001, *Social mobility: A discussion paper*, Performance and Innovation Unit, the Cabinet Office.

Badel, A. and Maues, J. S.: 2014, Why is economic mobility in Memphis among the lowest in the nation?, *The Regional Economist* **22**(2), 1–2.

Basile, R. and Gress, B.: 2005, Semi-Parametric Spatial Auto-Covariance Models of Regional Growth in Europe, *Region et Developpement* **21**, 93–118.

Becker, G. S.: 1991, *A Treatise on the Family*, Harvard University Press, Cambridge, Mass.

Belitz, C., Brezger, A., Klein, N., Kneib, T., Lang, S. and Umlauf, N.: 2014, BayesX - Software for Bayesian inference in structured additive regression models. Version 2.2.
URL: *http://www.bayesx.org*

Bengali, L. and Daly, M.: 2013, US economic mobility: the dream and the data, *Federal Reserve Bank of San Francisco Economic Letter* (06), 1–5.

Black, S. E. and Devereux, P. J.: 2010, Recent developments in intergenerational mobility, *in* O. Ashenfelter and D. Card (eds), *Handbook of labor economics, Volume 4b*, North Holland, pp. 1487–1541.

Brady, H. E., Schlozman, K. L. and Verba, S.: 2015, Political Mobility and Political Reproduction from Generation to Generation, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 149–173.

Brezger, A. and Lang, S.: 2006, Generalized structured additive regression based on Bayesian P-splines, *Computational Statistics & Data Analysis* **50**(4), 967–991.

Chetty, R., Hendren, N., Kline, P. and Saez, E.: 2013, The Economic Impacts of Tax Expenditures: Evidence from Spatial Variation Across the U.S., *Working Paper* pp. 1–33.

Chetty, R., Hendren, N., Kline, P. and Saez, E.: 2014, Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States, *The Quarterly Journal of Economics* **129**(4), 1553–1623.

Corak, M.: 2013, Income inequality, equality of opportunity, and intergenerational mobility, *The Journal of Economic Perspectives* **27**(3), 79–102.

Downs, G. W. and Rocke, D. M.: 1979, Interpreting heteroscedasticity, *American Journal of Political Science* **23**(4), 816–828.

Duncan, B. and Trejo, S. J.: 2015, Assessing the Socioeconomic Mobility and Integration of U.S. Immigrants and Their Descendants, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 108–135.

Dunn, P. K. and Smyth, G. K.: 1996, Randomized quantile residuals, *Journal of Computational and Graphical Statistics* **5**(3), 236–244.

Eilers, P. H. C. and Marx, B. D.: 1996, Flexible smoothing with B-splines and penalties, *Statistical science* **11**(2), 89–121.

Fahrmeir, L. and Kneib, T.: 2011, *Bayesian smoothing and regression for longitudinal, spatial and event history data*, Vol. 36 of *Oxford statistical science series*, Oxford University Press, Oxford and New York.

Fahrmeir, L., Kneib, T., Lang, S. and Marx, B.: 2013, *Regression: Models, methods and applications*, 1 edn, Springer, Berlin.

Fields, G. S. and Ok, E. A.: 1999, *The measurement of income mobility: an introduction to the literature*, Springer.

Fotheringham, A. S., Brunsdon, C. and Charlton, M.: 2002, *Geographically weighted regression: The analysis of spatially varying relationships*, Wiley, Chichester and West Sussex and Hoboken and NJ.

Genicot, G. and Ray, D.: 2012, Measuring Upward Mobility, *New York: Russell Sage Working Paper* pp. 1–11.

Goodman, L. A.: 1953, Ecological regressions and behavior of individuals, *American Sociological Review* **18**(6), 663–664.

Goodman, L. A.: 1959, Some alternatives to ecological correlation, *American Journal of Sociology* **64**(6), 610–625.

Grawe, N. D. and Mulligan, C. B.: 2002, Economic interpretations of intergenerational correlations, *Journal of Economic Perspectives* **16**(3), 45–58.

Grusky, D. B., Smeeding, T. M. and Snipp, C. M.: 2015, A New Infrastructure for Monitoring Social Mobility in the United States, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 63–82.

Hannan, M. T. and Burstein, L.: 1974, Estimation from grouped observations, *American Sociological Review* **39**(3), 374–392.

Haskins, R., Isaacs, J. and Sawhill, I.: 2008, Getting ahead or losing ground: Economic mobility in America, *Washington: The Brookings Institution* pp. 1–120.

Hofert, M., Kojadinovic, I., Maechler, M. and Yan, J.: 2014, copula: Multivariate Dependence with Copulas.
URL: *http://CRAN.R-project.org/package=copula*

Hofert, M. and Mächler, M.: 2011, Nested Archimedean Copulas Meet R: The nacopula Package, *Journal of Statistical Software* **39**(9), 1–20.
URL: *http://www.jstatsoft.org/v39/i09/*

Hout, M.: 2015, A Summary of What We Know about Social Mobility, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 27–36.

Irwin, L. and Lichtman, A. J.: 1976, Across the great divide: Inferring individual level behavior from aggregate data, *Political Methodology* **3**(4), 411–439.

Iversen, G. R.: 1973, Recovering individual data in the presence of group and individual effects, *American Journal of Sociology* **79**(2), 420–434.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B. and Michelson, S.: 1972, *Inequality: A reassessment of the effect of family and schooling in America*, Harper, New York.

Joe, H.: 2014, *Dependence Modeling with Copulas*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Johnson, D. S., Massey, C. and O'Hara, A.: 2015, The Opportunities and Challenges of Using Administrative Data Linkages to Evaluate Mobility, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 247–264.

Juhn, C., Murphy, K. M. and Pierce, B.: 1993, Wage inequality and the rise in returns to skil, *Journal of Political Economy* **101**(3), 410–442.

Kauermann, G., Haupt, H. and Kaufmann, N.: 2012, A hitchhiker's view on spatial statistics and spatial econometrics for lattice data, *Statistical Modelling* **12**(5), 419–440.

King, G.: 2013, *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data*, Princeton University Press.

Klein, N., Kneib, T., Klasen, S. and Lang, S.: 2015, Bayesian structured additive distributional regression for multivariate responses, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **64**(4), 569–591.

Klein, N., Kneib, T. and Lang, S.: 2013, Bayesian Structured Additive Distributional Regression, *Technical Report. Georg-August-Universität Göttingen* .
URL: *http://eeecon.uibk.ac.at/wopec2/repec/inn/wpaper/2013-23.pdf*

Kleinbaum, D. G. and Kupper, L. L.: 1982, *Epidemiologic research: principles and quantitative methods*, John Wiley & Sons.

Kojadinovic, I. and Yan, J.: 2010, Modeling Multivariate Distributions with Continuous Margins Using the copula R Package, *Journal of Statistical Software* **34**(9), 1–20.
URL: *http://www.jstatsoft.org/v34/i09/*

Kostov, P.: 2009, Spatial dependence in agricultural land prices: does it exist?, *Agricultural Economics* **40**(3), 347–353.

Kourtellos, A., Marr, C. and Tan, C. M.: 2015, Robust Determinants of Intergenerational Mobility in the Land of Opportunity, *European Economic Review (2015)* .
URL: *http://dx.doi.org/10.1016/j.euroecorev.2015.07.008i*

Krueger, A.: 2012, The Rise and Consequences of Inequality in the United States: Speech at the Center for American Progress, Washington D.C. on January 12, 2012.

Lang, S., Adebayo, S. B., Fahrmeir, L. and Steiner, W. J.: 2003, Bayesian geoadditive seemingly unrelated regression, *Computational Statistics & Data Analysis* **18**(2), 263–292.

Lang, S. and Brezger, A.: 2004, Bayesian P-splines, *Journal of Computational and Graphical Statistics* **13**(1), 183–212.

Lemieux, T.: 2006, Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?, *The American Economic Review* **96**(3), 461–498.

Mare, R. D.: 2015, Measuring Networks beyond the Origin Family, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 97–107.

Mazumder, B. and Acosta, M.: 2015, Using Occupation to Measure Intergenerational Mobility, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 174–193.

McMillen, D. P.: 2003, Spatial autocorrelation or model misspecification?, *International Regional Science Review* **26**(2), 208–217.

Meckstroth, T. W.: 1974, Some Problems in Cross-Level Inference, *American Journal of Political Science* **18**(1), 45–66.

Muller, C.: 2015, Measuring Education and Skil, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 136–148.

Murray, C. A.: 1984, *Losing ground: American social policy, 1950-1980*, Basic books, New York.

Murray, C. A.: 2012, *Coming apart: The state of white America, 1960-2010*, N.Y.: Crown Forum, New York.

Narey, M.: 2009, Report from the Independent Commission on Social Mobility, *London, Liberal Democrats* pp. 1–85.

Nelder, J. A. and Lee, Y.: 1991, Generalized linear models for the analysis of taguchi–type experiments, *Applied stochastic models and data analysis* **7**(1), 107–120.

Nelsen, R. B.: 2007, *An introduction to copulas*, Springer.

Nickerson, C. and Zhang, W.: 2014, Modeling the Determinants of Farmland Values in the United States, *in* J. Duke and J. Wu (eds), *The Oxford Handbook of Land Economics*, Oxford: Oxford University Press, pp. 111–137.

Piketty, T. and Saez, E.: 2003, Income Inequality in the United States, 1913-1998, *The Quarterly Journal of Economics* **118**(1), 1–41.

Prewitt, K.: 2015, Who Is Listening? When Scholars Think They Are Talking to Congress, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 265–273.

Reeves, R.: 2015, The Measure of a Nation, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 22–26.

Reeves, R. V. and Sawhill, I. V.: 2014, Equality of Opportunity: Definitions, Trends, and Interventions, *Federal Reserve Bank of Boston Working Paper* pp. 1–55.

Robinson, W. S.: 1950, Ecological Correlations and the Behavior of Individuals, *American Sociological Review* **15**(3), 351–357.

Selvin, H. C.: 1965, Durkheim's suicide: Further thoughts on a methodological classic, *in* R. A. Nisbet (ed.), *Émile Durkheim*, pp. 113–137.

Sharkey, P. and Graham, B.: 2013, Mobility and the Metropolis: How Communities Factor Into Economic Mobility, *The Pew Charitable Trusts* pp. 1–23.

Smyth, G. K.: 1989, Generalized linear models with varying dispersion, *Journal of the Royal Statistical Society. Series B (Methodological)* **51**(1), 47–60.

Smyth, G. K.: 2002, An efficient algorithm for REML in heteroscedastic regression, *Journal of Computational and Graphical Statistics* **11**(4), 836–847.

Smyth, G. K., Huele, A. F. and Verbyla, A. P.: 2001, Exact and approximate REML for heteroscedastic regression, *Statistical Modelling* **1**(3), 161–175.

Snellman, K., Silva, J. M., Frederick, C. B. and Putnam, R. D.: 2015, The Engagement Gap: Social Mobility and Extracurricular Participation among American Youth, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 194–207.

Solon, G.: 1965, Intergenerational mobility in the labor market, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics, Volume 3, Part A*, pp. 1761–1800.

Solon, G.: 1992, Intergenerational income mobility in the United States, *The American Economic Review* **82**(3), 393–408.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A.: 2002, Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(4), 583–639.

Tach, L.: 2015, Social Mobility in an Era of Family Instability and Complexity, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 83–96.

The Pew Charitable Trusts: 2009, Opinion Poll on Economic Mobility and the American Dream.
URL: *http://www.pewtrusts.org/en/multimedia/data-visualizations/2013/income-and-wealth-in-america-across-generations*

Torche, F.: 2015, Analyses of Intergenerational Mobility: An Interdisciplinary Review, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 37–62.

Urahn, S. K., Currier, E., Elliott, D., Wechsler, L., Wilson, D. and Colbert, D.: 2012, Pursuing the American dream: Economic mobility across generations, *The Pew Charitable Trusts* pp. 1–38.

Warren, J. R.: 2015, Potential Data Sources for a New Study of Social Mobility in the United States, *The ANNALS of the American Academy of Political and Social Science* **657**(1), 208–246.

Western, B. and Bloome, D.: 2009, Variance function regressions for studying inequality, *Sociological Methodology* **39**(1), 293–326.

Western, B., Bloome, D. and Percheski, C.: 2008, Inequality among American families with children, 1975 to 2005, *American Sociological Review* **73**(6), 903–920.

Wilson, W. J.: 1996, *When work disappears: The world of the new urban poor*, Knopf: Distributed by Random House, Inc., New York.

Yan, J. : 2007, Enjoy the Joy of Copulas: With a Package copula, *Journal of Statistical Software* **21**(4), 1–21.
URL: *http://www.jstatsoft.org/v21/i04/*

Zellner, A.: 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association* **57**(298), 348–368.

Zheng, H., Yang, Y. and Land, K. C.: 2011, Variance Function Regression in Hierarchical Age-Period-Cohort Models: Applications to the Study of Self-Reported Health, *American Sociological Review* **76**(6), 955–983.

Zheng, H., Yang, Y. and Land, K. C.: 2013, Heteroscedastic Regression Models for the Systematic Analysis of Residual Variances, *in* S. L. Morgan (ed.), *Handbook of Causal Analysis for Social Research*, Springer, Dordrecht, pp. 133–152.

Notes

[1]This effect is usually referred to as 'stickiness at the ends' (Haskins et al., 2008).

[2]See also Genicot and Ray (2012) that combine aspects of both types of mobility into a single measure.

[3]One might argue that the availability of economic theory or a-priori knowledge makes an analysis of heteroscedasticity largely unnecessary. While it is true that there are examples where substantive knowledge and theoretical understanding provide some insight into heteroscedasticity (e.g., in a regression of consumption on income, where in comparison to high incomes, the consumption level of low income people is bounded on the low side, largely due to necessary expenditures for survival that already consume a large portion of peoples income, or due to differential access to credit and accumulated assets of low income people), there are, however, frequent cases in which such knowledge has to be derived directly from the data, so that the possibility that heteroscedasticity can lead to new insight and theory building suggests that it should be analyzed (Downs and Rocke, 1979).

[4]Although the framework of Bayesian structured additive multivariate distributional regression of Klein et al. (2015) can in general be extended to the D-dimensional case, we restrict our presentation to the bivariate case only, where D=2.

[5]For the bivariate Normal distribution, the density of $\boldsymbol{y} = (y_1, y_2)'$ is defined as follows:

$$p(y_1, y_2) = \frac{1}{\sqrt{(2\pi)^2|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right)$$

with expectation $\boldsymbol{\mu} = (E(y_1), E(y_2))'$ and positive semidefinite covariance matrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$.

For the bivariate Student-t distribution, the density of $\boldsymbol{y} = (y_1, y_2)'$ takes the following form:

$$p(y_1, y_2) = \frac{\Gamma\left((n_{df}+2)/2\right)}{\Gamma(n_{df}/2)(n_{df}\pi)}|\boldsymbol{\Sigma}|^{-1/2}\left(1 + (\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right)^{-(n_{df}+2)/2}.$$

with location parameters $\boldsymbol{\mu} = (\mu_1, \mu_2)'$, dispersion matrix $\boldsymbol{\Sigma}$ and degrees of freedom $n_{df} > 0$.

[6]Functional form misspecifications of the estimated effects can also adversely influence statistical inference, as they may affect both the magnitude and the significance of the estimated effects. In particular, misspecifications with respect to the functional form may be responsible for finding spurious spatial dependencies. For details see Nickerson and Zhang (2014); Kostov (2009); Basile and Gress (2005); McMillen (2003). In order to alleviate such problems, we relax the linearity assumption and allow for a data-driven selection of the functional form.

[7]Since Bayesian distributional regression and the estimation of the spatial effects are motivated from a statistical point of view and since many readers may be more familiar with models from

spatial econometrics, we refer the interested reader to Kauermann et al. (2012) who contrasts and highlights the differences between models from spatial statistics and spatial econometrics.

[8]The data is available online and can be downloaded from the EQUALITY OF OPPORTUNITY PROJECT-website at www.equality-of-opportunity.org. All calculations in this paper are based on Version 2.0 of the data, released January 17, 2014.

[9]Commuting Zones are aggregations of counties, where each commuting zone consists on average of 4 counties with an average population of 380,000.

[10]The 25-*th* percentile of the national parent family income distribution yields an equivalent of $28,800.

[11]Since Chetty et al. (2014) compare children to those in the nation as a whole rather than to those in their own county or commuting zone, the expected mean income rank is in fact a measure of absolute mobility. In particular, Chetty et al. (2014) consider $\bar{r}_{25c}$ as a way of answering the question 'What are the outcomes of children from families who grow up in low-income families?' Consequently, the fact that Chetty et al. (2014) rank both children and parents in a certain commuting zone based on their positions in the national income distribution, allows the authors to measure children's absolute outcomes using $\bar{r}_{25c}$ as their preferred measure of absolute mobility. Moreover, Chetty et al. (2014) argue that 'at the national level, the measure of absolute mobility is mechanically related to the rank-rank slope and does not provide any additional information over the other measure'. However, Chetty et al. (2014) continue by saying that 'when we study small areas within the U.S., a child's rank in the national income distribution is effectively an absolute outcome because incomes in a given area have little impact on the national distribution'.

[12]As shown by Chetty et al. (2014) and Kourtellos et al. (2015), the variables in the parent group 'College education' do not seem to influence either absolute or relative mobility so that the exclusion of these variable should not alter the analysis.

[13]The estimated effects are obtained via full Bayesian MCMC simulation based on 320,000 iterations, a burn-in period of 20,000 iterations and a thinning parameter of 300 resulting in a sample of 1,000 samples from the posterior. For all estimated effects, we investigated the resulting sampling paths for convergence and mixing and found no evidence of remaining autocorrelation. The smooth effects of continuous covariates are estimated via cubic penalized B-splines based on second order random walk priors and 10 equidistant inner knots for all model specifications. Hyper-parameters for the smoothing variances τ_j^2 are set to $a_j = b_j = 0.001$ as a default, as these parameter values have turned out to be a robust choice in previous studies. To analyze the sensitivity of the estimated effects with respect to the choice of the hyper-parameters, we re-estimated the models for values of $a_j = 1.0, b_j = 0.005$ and found no differences in the results. The structured spatial effect is estimated based on a Markov random field prior. For the stepwise forward selection, we reduced the number of iterations to 12,000 to make the variable selection computationally feasible. Both variable selection and the estimation are performed using version 2.2 of the software *BayesX* of Belitz et al. (2014).

[14]We have also tried to fit the Gumbel copula to the data. However, we experienced numerical

problems. Besides being an indication that the Gumbel is not appropriate for the data, the numerical problems might be related to the fact that the Gumbel copula does not allow for negative dependence.

[15]Estimation was performed using the R-package *copula* of Kojadinovic, I. and Yan, J. (2010); Yan, J. (2007); Hofert, M. et al. (2014); Hofert, M. and Mächler, M. (2011).

[16]For more details on the ecological fallacy and discussions thereof see Goodman (1953), Goodman (1959), Iversen (1973), Hannan and Burstein (1974), Meckstroth (1974), Irwin and Lichtman (1976), Kleinbaum and Kupper (1982) and King (2013).

[17]As the model used for the analysis allows for the explicit modeling of such unobserved heterogeneity, the spatial patterns presented in our analysis seem to support this conjecture, as many of the spatial patterns remain unexplained, especially for the mean regression models.

[18]Since the estimated effects for $\eta^{n_{df}}$ are difficult to interpret economically and since our hypotheses are concerned with the first five moments of the Student-t distribution, we restrict our analysis to $\eta^{\mu_{abs}}$, $\eta^{\mu_{rel}}$, $\eta^{\sigma^2_{abs}}$, $\eta^{\sigma^2_{rel}}$ and η^{ρ}.

[19]Although Chetty et al. (2014) present the results on the commuting zone level, we restrict the presentation of the spatial effects to the level of the 432 US-congressional districts in the present analysis. This change of the spatial scale is due to the low mixing behaviour of the Markov Chains when commuting zones are used, especially for the spatial effects of the higher moments of the response distributions. Mixing problems may be attributed to several factors: First, the possibility that a spatial effect is selected for each of the K distributional parameters, as well as the fact that each continuous covariate effect is modeled as a smooth non-linear function leads to a problem that is usually encountered with high dimensional data in the sense that the number of parameters to be estimated might exceed the number of observations, even though the equivalent degrees of freedom are reduced due to penalization. Second, the fact that, on average, there are only 4 counties within each commuting zone might pose an additional problem for the estimation of the spatial effects, as the informational content from each commuting zone might be too little in order to clearly identify and separate between spatial effects for the higher moments. In our efforts to eliminate this problem, the average number of counties doubles to eight so that the mixing behaviour of the Markov Chains improve as we move from commuting zones to congressional districts. For ease of visualization and interpretation, the spatial effects are presented at the district-level. The effects of the covariate effects, however, are presented at the county-level.

[20]All estimated effects are found in the Tables and Figures section.

[21]One reason why our results differ from Kourtellos et al. (2015) may be due to the fact, that Kourtellos et al. (2015) assume a linear functional form between the different measures of social mobility while we allow for possible non-linearities in the data by modeling the covariate effects in a non-linear way using P-splines. The differing results might also be attributed to the slightly different data set that we use for our analysis compared to that Chetty et al. (2014) and Kourtellos et al. (2015).

[22]See Chetty et al. (2013) for an overview of the effects of tax expenditures on intergenerational

social mobility.

[23]Figures 3, 4, 6 and 7 show posterior mean estimates for the semi-parametric effects together with 95% pointwise credible intervals (shaded area in grey). Since the parametric part of the additive predictor includes an overall intercept term, each function is centered around zero, i.e., $\sum_{i=1}^{n} f_1(z_{i1}) = \ldots = \sum_{i=1}^{n} f_p(z_{ip}) = 0$ in order to guarantee the identification of the estimated semi-parametric effects. The estimated functions in Figures 3, 4, 6 and 7 are plotted in this way. The x-axis shows the domain of the corresponding centered covariate, as well as a rug plot indicating the positions of the data points along the x-axis denoted by tick marks. With respect to the general direction of the estimated effects, Figures 3 and 4 show that they are generally in accordance with the results presented in Chetty et al. (2014) and Kourtellos et al. (2015).

[24]The same interpretation holds true in the case of a continuous covariate, where groups are defined as individual observations defined over the domain of the corresponding covariate.

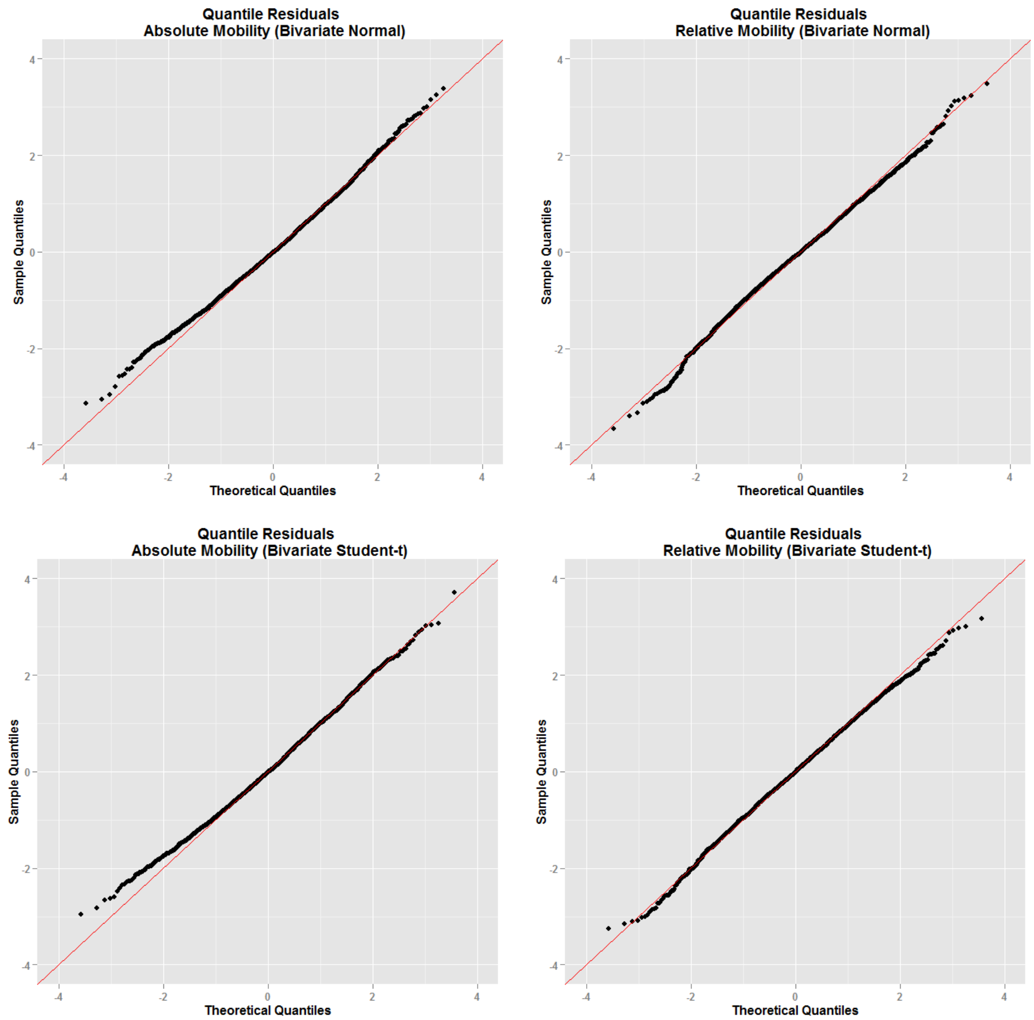
[25]The Great Gatsby Curve suggests that levels of income inequality and levels of social mobility are negatively correlated so that a rise in income inequality for the current generation of families could lead to a slowdown in social mobility for the next generation. See Krueger (2012) and Corak (2013) for details.

[26]Making use of the fact that both measures are correlated also allows for a more precise prediction of future rates of social mobility.

[27]In contrast to the estimated effects of all other distributional parameters, the effects shown in Figure 10 are not centered around zero. Instead, Figure 10 shows the posterior mean estimates of the semi-parametric effects of each covariate effect on ρ transformed to the interval $[-1, 1]$ together with 95% pointwise credible intervals, with the transformation $\eta^{\rho}/\sqrt{1 + (\eta^{\rho})^2}$, where $\eta^{\rho} = \beta_0^{\rho} + f^{\rho}(\cdot)$. All other effects are set to 0. The dashed line indicates the unweighted coefficient of correlation of -0.68 between the two measures as reported by Chetty et al. (2014).

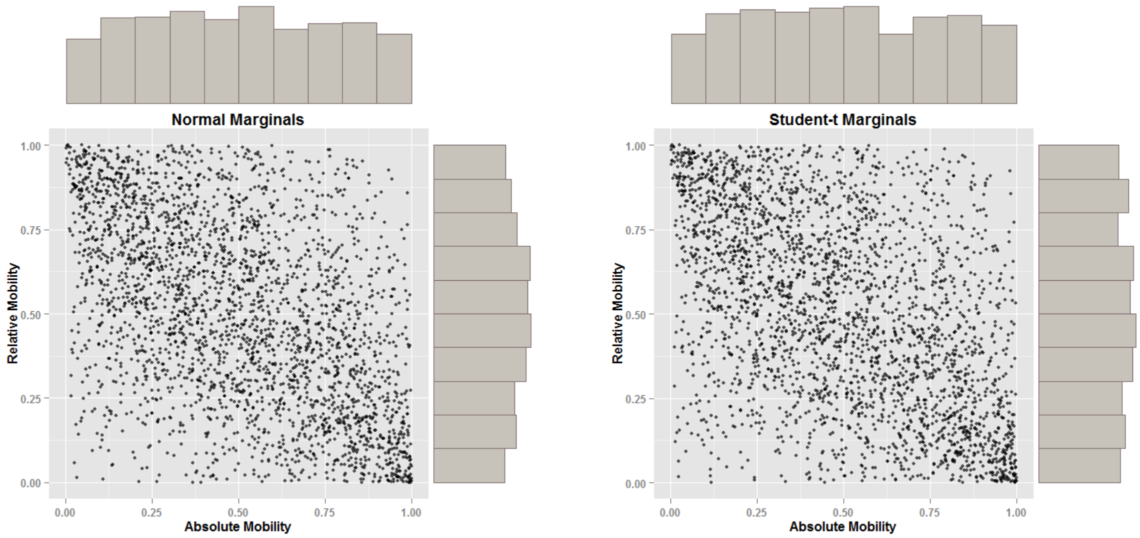
[28]For an overview of the literature, see Solon (1965), Grawe and Mulligan (2002) or Black and Devereux (2010). The increasing attention that social mobility has received in recent years, not only in political, but also in scholarly circles is represented in a recent special issue of *The ANNALS of the American Academy of Political and Social Science*, where an overview of the topic is provided by leading scholars in the field. In particular see Reeves (2015), Hout (2015), Torche (2015), Grusky et al. (2015), Tach (2015), Mare (2015) Duncan and Trejo (2015), Muller (2015), Brady et al. (2015), Mazumder and Acosta (2015), Snellman et al. (2015), Warren (2015), Johnson et al. (2015) and Prewitt (2015). The recent interest in social mobility has also fueled by Krueger (2012) and Corak (2013) who find a negative relationship between levels of income inequality and levels of social mobility across high-income countries (also known as the 'Great Gatsby Curve').

Figures



Source: Author's calculations based on data from Chetty et al. (2014).

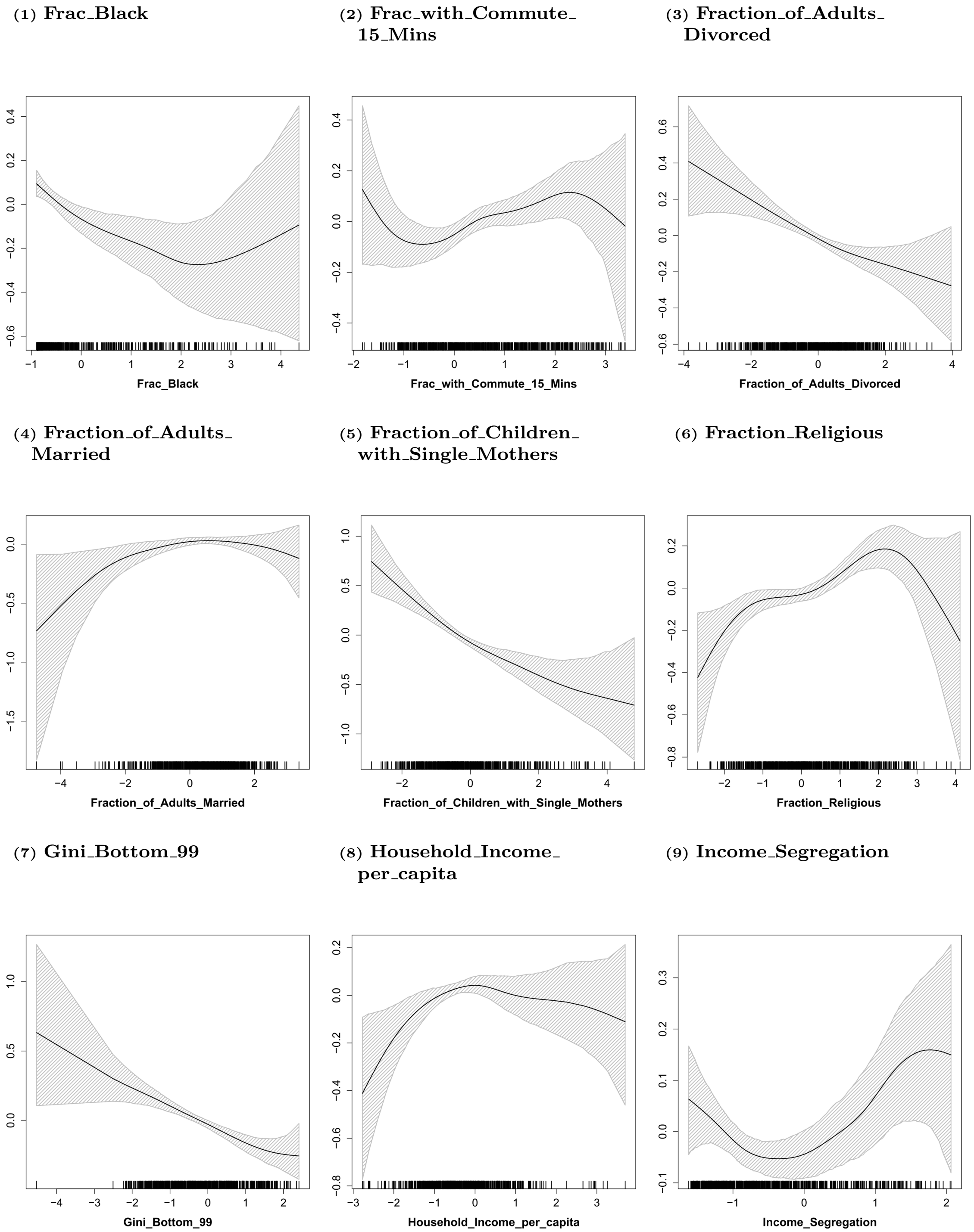
Figure 1
Quantile residuals of the marginals for the bivariate Normal and the bivariate Student-t distribution.



Source: Author's calculations based on data from Chetty et al. (2014).

Figure 2

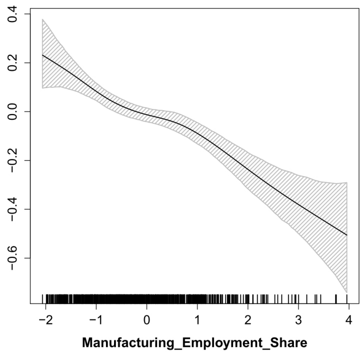
Uniform marginals of transformed absolute ($\bar{r}_{25c}$) and relative mobility (β_c).



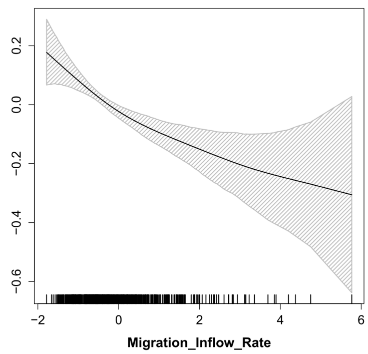
Source: Author's calculations based on data from Chetty et al. (2014).

Figure 3
Semi-parametric effects of mean absolute mobility ($\bar{r}_{25c}$) with 95% pointwise credible intervals (shaded area in grey).

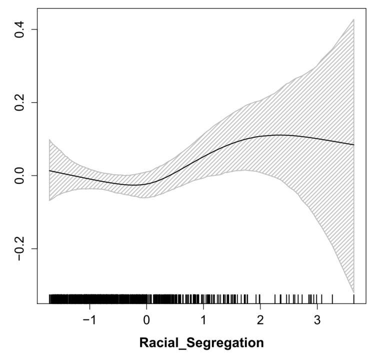
(10) Manufacturing_Employment_Share



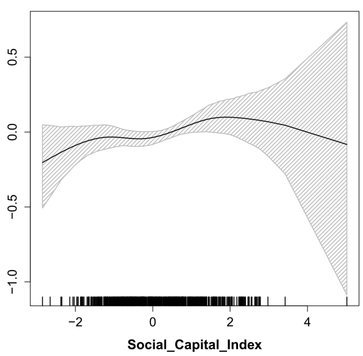
(11) Migration_Inflow_Rate



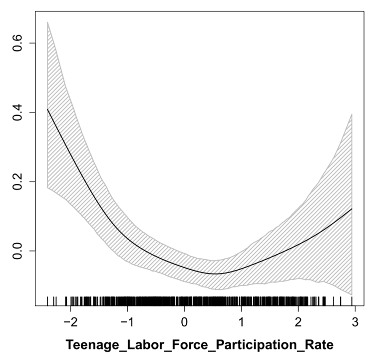
(12) Racial_Segregation



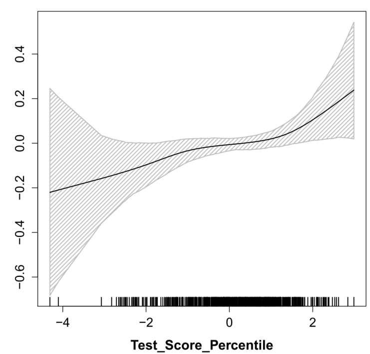
(13) Social_Capital_Index



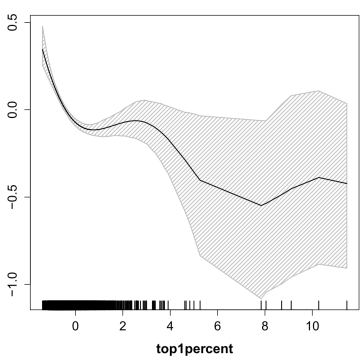
(14) Teenage_Labor_Force_Participation_Rate



(15) Test_Score_Percentile

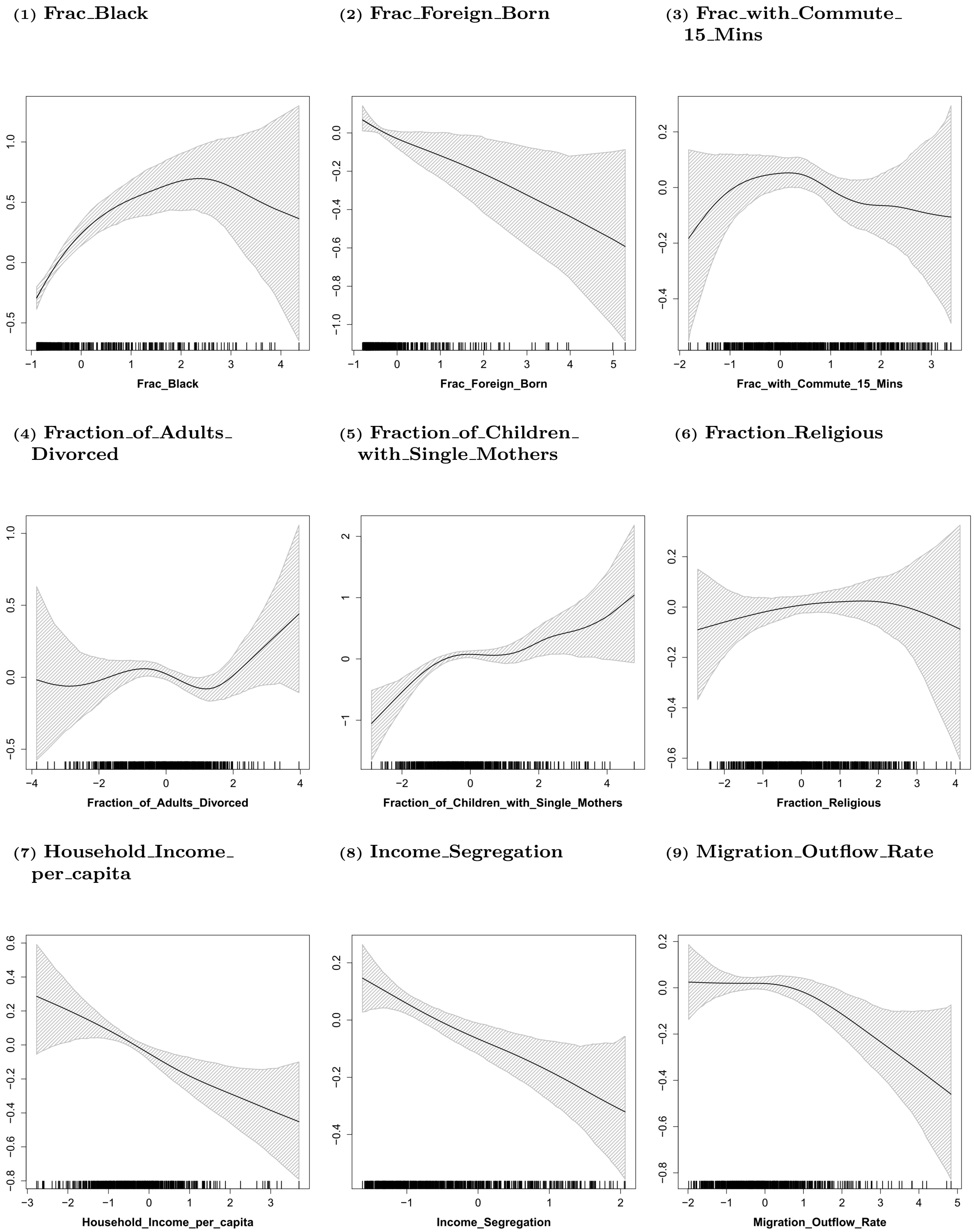


(16) top1percent



Source: Author's calculations based on data from Chetty et al. (2014).

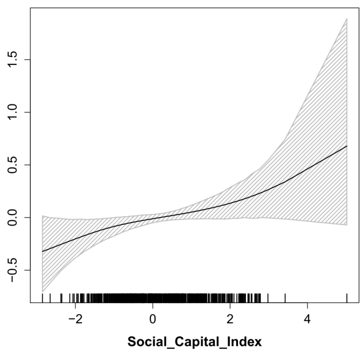
Figure 3
continued.



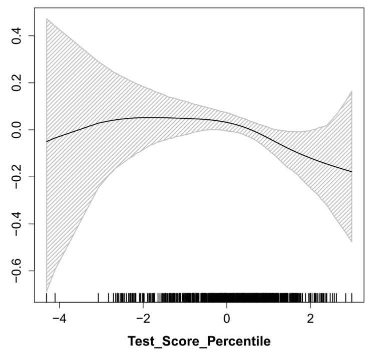
Source: Author's calculations based on data from Chetty et al. (2014).

Figure 4
Semi-parametric effects of mean relative mobility (β_c) with 95% pointwise credible intervals (shaded area in grey).

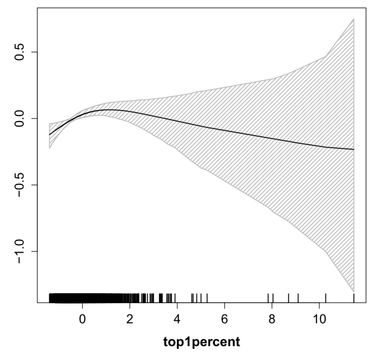
(10) Social_Capital_Index



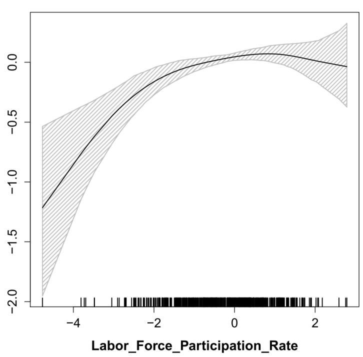
(11) Test_Score_Percentile



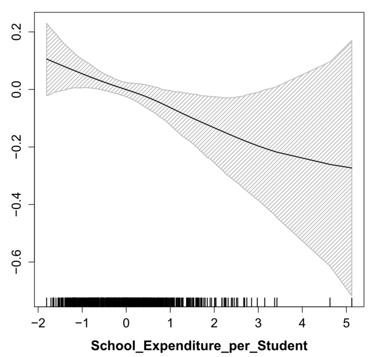
(12) top1percent



(13) Labor_Force_ Participation_Rate

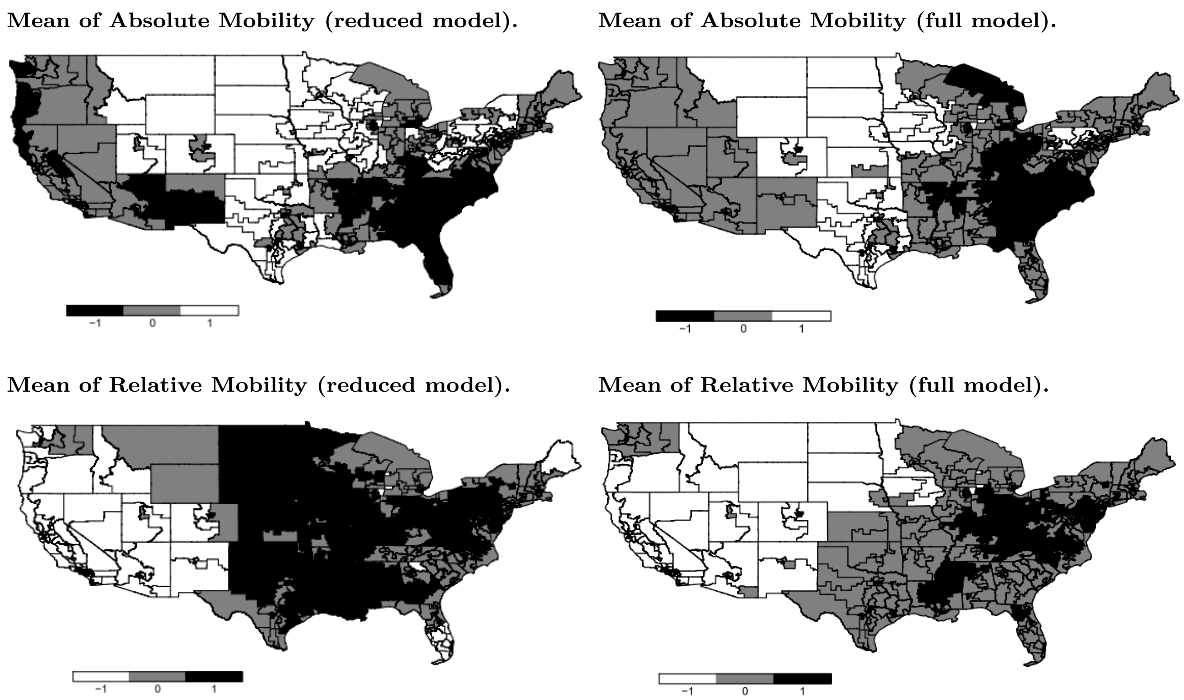


(14) School_Expenditure_ per_Student



Source: Author's calculations based on data from Chetty et al. (2014).

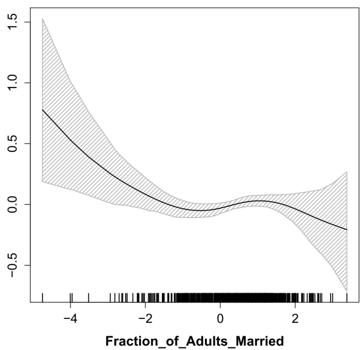
Figure 4
continued.



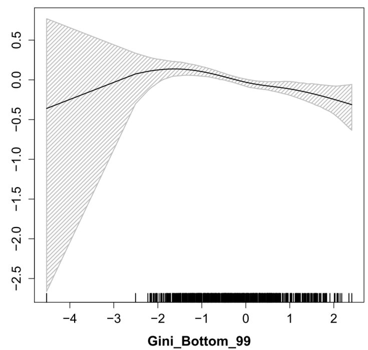
Source: Author's calculations based on data from Chetty et al. (2014).
-1: significantly negative; 0: non-significant; 1: significantly positive.

Figure 5
Significance of the spatial effects of mean absolute ($\bar{r}_{25c}$) and relative mobility (β_c) based on posterior probabilities using a nominal level of 95%.

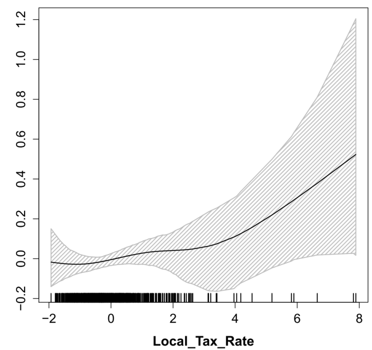
(1) **Fraction_of_Adults_Married**



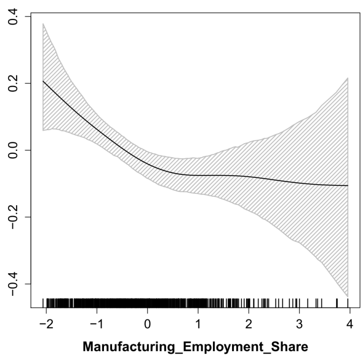
(2) **Gini_Bottom_99**



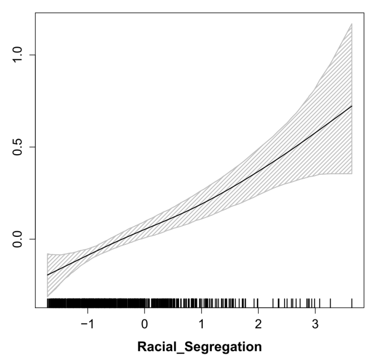
(3) **Local_Tax_Rate**



(4) **Manufacturing_Employment_Share**



(5) **Racial_Segregation**

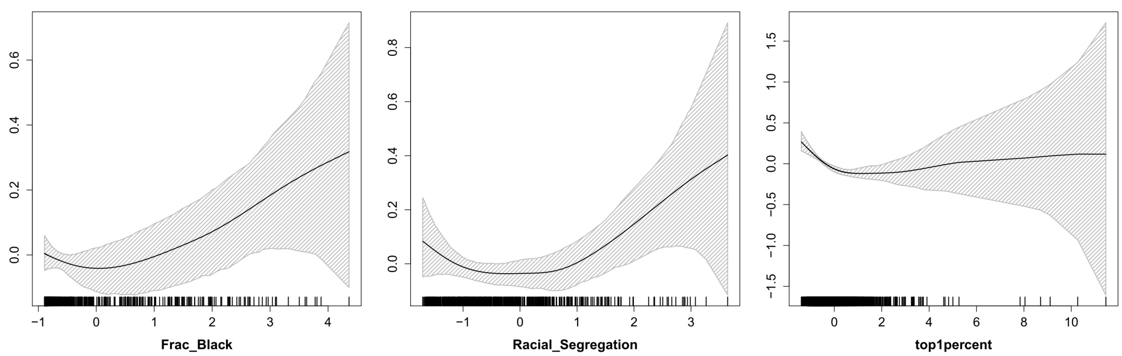


Source: Author's calculations based on data from Chetty et al. (2014).

Figure 6

Semi-parametric effects of within-district variance of absolute mobility ($\bar{r}_{25c}$) with 95% pointwise credible intervals (shaded area in grey).

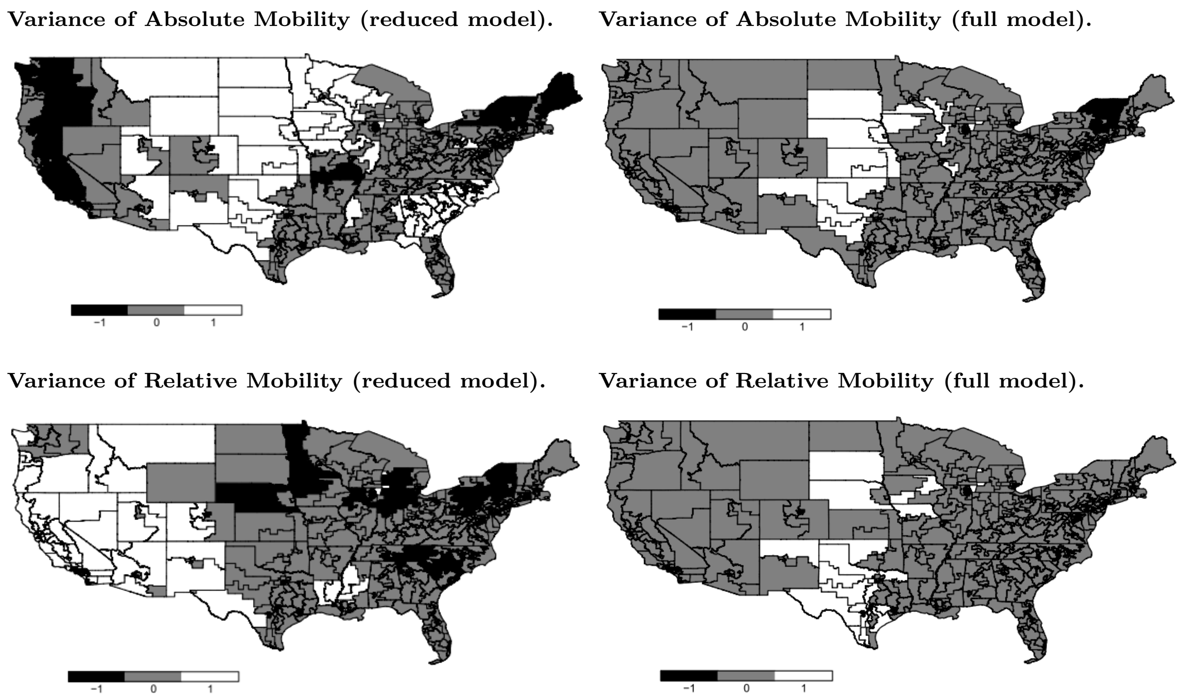
(1) **Frac_Black** (2) **Racial_Segregation** (3) **top1percent**



Source: Author's calculations based on data from Chetty et al. (2014).

Figure 7

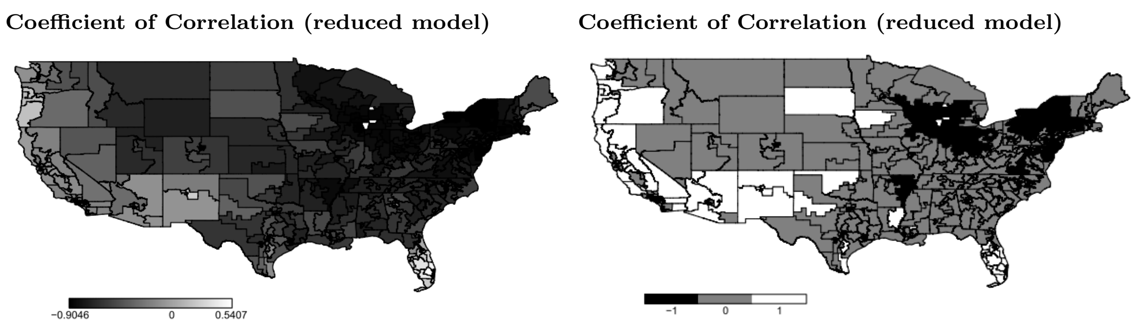
Semi-parametric effects of within-district variance of relative mobility (β_c) with 95% pointwise credible intervals (shaded area in grey).



Source: Author's calculations based on data from Chetty et al. (2014).
-1: significantly negative; 0: non-significant; 1: significantly positive.

Figure 8

Significance of the spatial effects of the within-district variance of absolute ($\bar{r}_{25c}$) and relative mobility (β_c) based on posterior probabilities using a nominal level of 95%.

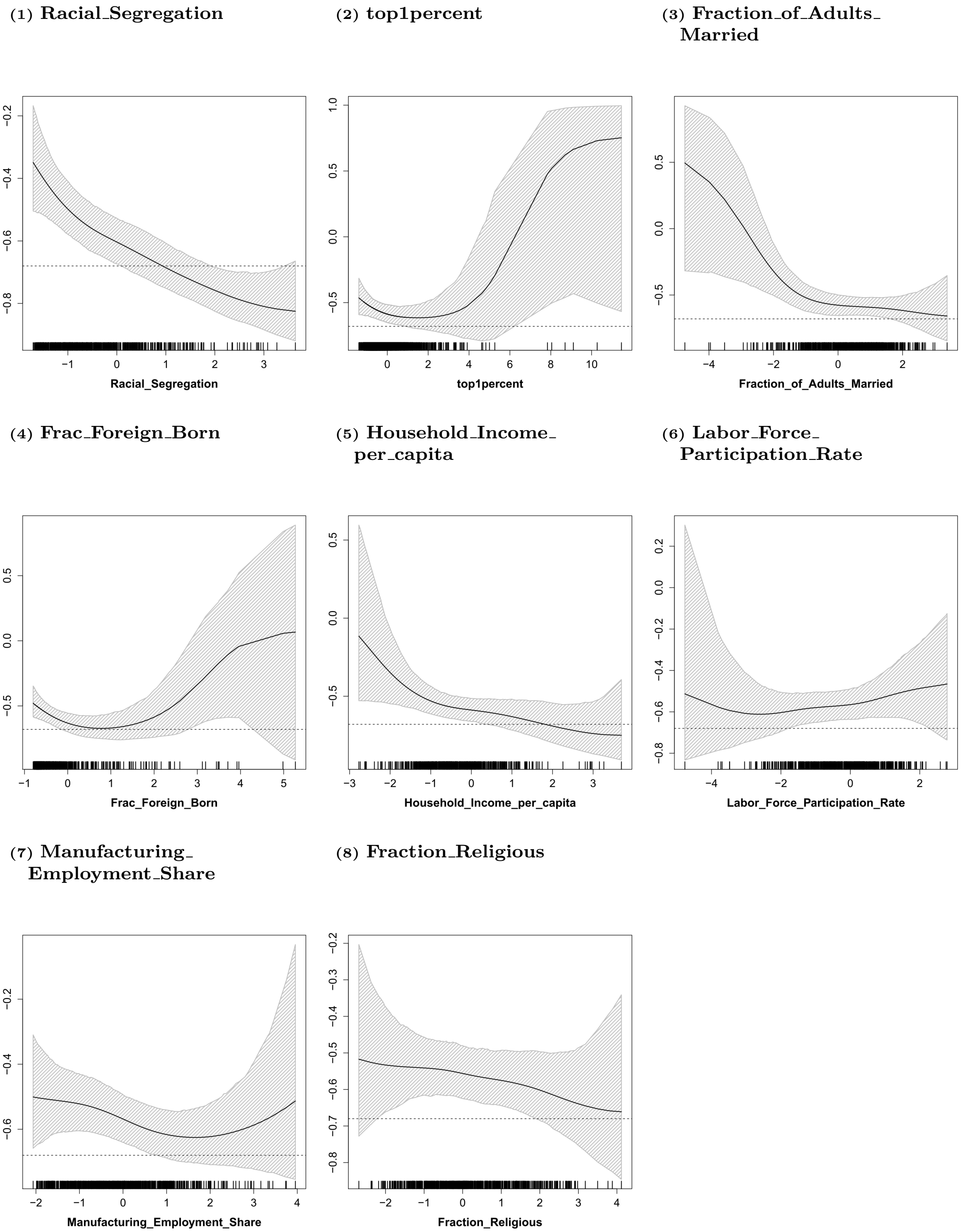


Source: Author's calculations based on data from Chetty et al. (2014).
-1: significantly negative; 0: non-significant; 1: significantly positive.

Figure 9

Posterior mean estimates (left panel) and significance of the spatial effect of ρ between absolute ($\bar{r}_{25c}$) and relative mobility (β_c) based on a nominal level of 95% (right panel).

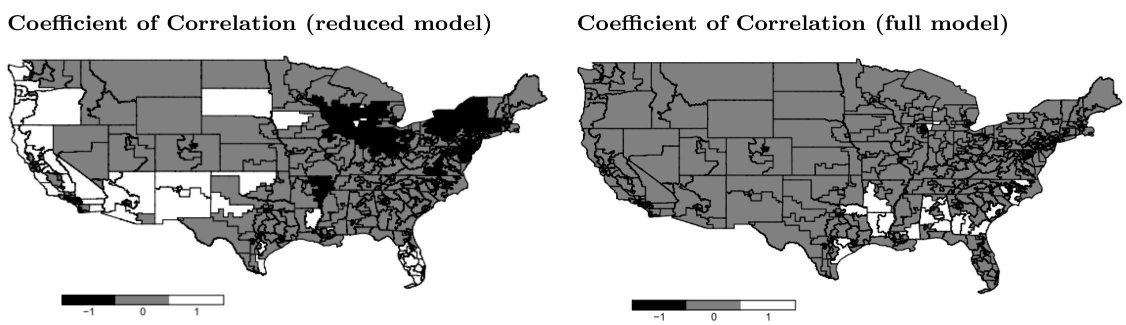
For the left panel, posterior mean estimates of the spatial effects are converted to the scale of the coefficient of correlation ρ of $[-1, 1]$ with the transformation $\eta^{\rho}_{geo}/\sqrt{1+(\eta^{\rho}_{geo})^2}$, where $\eta^{\rho}_{geo} = \beta^{\rho}_{geo,0} + f^{\rho}_{geo}$. All other covariate effects are set to 0.



Source: Author's calculations based on data from Chetty et al. (2014).

Figure 10
Semi-parametric effects of ρ between absolute ($\bar{r}_{25c}$) and relative mobility (β_c) with 95% pointwise credible intervals (shaded area in grey).

Shown are the posterior mean estimates of the semi-parametric effects of each covariate effect converted to the scale of the coefficient of correlation ρ of $[-1, 1]$ with the transformation $\eta^{\rho}/\sqrt{1+(\eta^{\rho})^2}$, where $\eta^{\rho} = \beta_0^{\rho} + f^{\rho}(\cdot)$. The dashed line indicates the unweighted coefficient of correlation of -0.68 between the two measures as reported by Chetty et al. (2014).



Source: Author's calculations based on data from Chetty et al. (2014).
-1: significantly negative; 0: non-significant; 1: significantly positive.

Figure 11
Significance of the spatial effects of ρ between absolute ($\bar{r}_{25c}$) and relative mobility (β_c) based on a nominal level of 95%.

Tables

Table 1
Description of Variables.

Variable (County Level)	Description	Parent Group
abs_mobility	Expected rank of children whose parents are at the 25th percentile of the national income distribution based on rank-rank regression (continuous).	
rel_mobility	Rank-rank slope from OLS regression of child rank on parent rank within each county in core sample using baseline income definitions (continuous).	
top1percent	Share of parent income within county accruing to the county's top 1 percent of tax filers for parents in core sample (continuous).	Income Distribution
Variable (Commuting Zone Level)	**Description**	
Gini_Bottom_99	Gini coefficient minus top 1% income share (continuous).	Income Distribution
Household_Income_per_capita	Aggregate household income in the 2000 census divided by the number of people aged 16-64 (continuous).	Income Distribution
Frac_Black	Number of individuals who are black alone divided by total population (continuous).	Segregation
Racial_Segregation	Multi-group Theil Index calculated at the census-tract level over four groups: White alone, Black alone, Hispanic, and Other (continuous).	Segregation
Income_Segregation	Rank-Order Theil index estimated at the census-tract level (continuous).	Segregation
Frac_with_Commute_15_Mins	Number of workers that commute less than 15 minutes to work divided by total number of workers (continuous).	Segregation
Local_Tax_Rate	Total tax revenue per capita divided by mean household income per capita for working age adults (in 2000) (continuous).	Tax
Local_Government_Expenditures_Per_Capita	Total local government expenditures per capita (continuous).	Tax
State_EITC_Exposure	The mean state EITC top-up rate between 1980-2001, with the rate coded as zero for states with no state EITC (continuous).	Tax
State_Income_Tax_Progressivity	The difference between the state income tax rate for incomes above $100,000 and incomes in the bottom tax income bracket (continuous).	Tax
School_Expenditure_per_Student	Average expenditures per student in public schools (continuous).	K-12 Education
Test_Score_Percentile	Residual from a regression of mean math and English standardized test scores on household income per capita in 2000 (continuous).	K-12 Education
Labor_Force_Participation_Rate	Share of people at least 16 years old that are in the labor force (continuous).	Local Labor Market
Manufacturing_Employment_Share	Share of employed persons 16 and older working in manufacturing (continuous).	Local Labor Market
Growth_in_Chinese_Imports_1990_2000	Percentage growth in imports from China per worker between 1990 and 2000, scaled as an annualized rate times 10 (continuous).	Local Labor Market
Teenage_Labor_Force_Participation_Rate	Fraction of children in birth cohorts 1985-1987 who received a W2 (i.e. had positive wage earnings) in any of the tax years when they were age 14-16 (continuous).	Local Labor Market
Migration_Inflow_Rate	Migration into the commuting zone from other commuting zones (divided by commuting zone population from 2000 Census) (continuous).	Migration
Migration_Outflow_Rate	Migration out of the commuting zone from other commuting zones (divided by commuting zone population from 2000 Census) (continuous).	Migration
Frac_Foreign_Born	Share of commuting zone residents born outside the United States (continuous).	Migration
Social_Capital_Index	Standardized index combining measures of voter turnout rates, the fraction of people who return their census forms, and measures of participation in community organizations (continuous).	Social Capital
Fraction_Religious	Share of religious adherents (continuous).	Social Capital
Fraction_of_Children_with_Single_Mothers	Number of single female households with children divided by total number of households with children (continuous).	Family Structure
Fraction_of_Adults_Divorced	Fraction of people 15 or older who are divorced (continuous).	Family Structure
Fraction_of_Adults_Married	Fraction of people 15 or older who are married and not separated (continuous).	Family Structure
Urban_Areas	Defined by Chetty et al. (2014) as commuting zones that intersect with metropolitan statistical areas (MSA) (binary).	

Source: Chetty et al. (2014).

Table 2
DIC values for the bivariate Normal and bivariate Student-t distribution.

Distribution	DIC
Bivariate Normal	8955.89
Bivariate Student-t	8756.55

Source: Author's calculations based on data from Chetty et al. (2014).

Table 3
Estimated Copulas and AIC.

	Normal Distribution	Student-t Distribution
Student-t Copula	-849.7846	-865.6534
Gaussian Copula	-815.0303	-815.5891
Frank Copula	-811.7049	-807.4675
Clayton Copula	-246.6960	-236.2641

Source: Author's calculations based on data from Chetty et al. (2014).

Table 4

Selected covariates and predictor specification for the bivariate Student-t distribution.

Predictor Specification
$\eta^{\mu_{abs}} = Intercept + Urban\_Areas + State\_Income\_Tax\_Progressivity + Growth\_in\_Chinese\_Imports\_1990\_2000 + f(fraction\_Religious) + f(fraction\_of\_Children\_with\_Single\_Mothers) + f(top1percent) + f(Migration\_Inflow\_Rate) + f(Manufacturing\_Employment\_Share) + f(frac\_with\_Commute\_15\_Mins) + f(frac\_Black) + f(Social\_Capital\_Index) + f(Household\_Income\_per\_capita) + f(Teenage\_Labor\_force\_Participation\_Rate) + f(Gini\_Bottom\_99) + f(fraction\_of\_Adults\_Divorced) + f(Income\_Segregation) + f(fraction\_of\_Adults\_Married) + f(Test\_Score\_Percentile) + f(Racial\_Segregation) + f_{geo}$
$\eta^{\mu_{rel}} = Intercept + State\_Income\_Tax\_Progressivity + Growth\_in\_Chinese\_Imports\_1990\_2000 + State\_EITC\_Exposure + f(Labor\_Force\_Participation\_Rate) + f(School\_Expenditure\_per\_Student) + f(Frac\_with\_Commute\_15\_Mins) + f(Fraction\_Religious) + f(Frac\_Black) + f(Fraction\_of\_Children\_with\_Single\_Mothers) + f(Household\_Income\_per\_capita) + f(Social\_Capital\_Index) + f(Migration\_Outflow\_Rate) + f(top1percent) + f(Test\_Score\_Percentile) + f(Fraction\_of\_Adults\_Divorced) + f(Income\_Segregation) + f(Frac\_Foreign\_Born) + f_{geo}$
$\eta^{\sigma^2_{abs}} = Intercept + f(Local\_Tax\_Rate) + f(Racial\_Segregation) + f(Manufacturing\_Employment\_Share) + f(Fraction\_of\_Adults\_Married) + f(Gini\_Bottom\_99) + f_{geo}$
$\eta^{\sigma^2_{rel}} = Intercept + Urban\_Areas + f(top1percent) + f(Racial\_Segregation) + f(Frac\_Black) + f_{geo}$
$\eta^{\rho} = Intercept + Urban\_Areas + State\_EITC\_Exposure + f(Racial\_Segregation) + f(top1percent) + f(Fraction\_of\_Adults\_Married) + f(Frac\_Foreign\_Born) + f(Manufacturing\_Employment\_Share) + f(fraction\_Religious) + f(Household\_Income\_per\_capita) + f(Labor\_Force\_Participation\_Rate) + f_{geo}$
$\eta^{n_{df}} = Intercept + Fraction\_of\_Children\_with\_Single\_Mothers + Racial\_Segregation + Household\_Income\_per\_capita$

Source: Author's calculation based on data from Chetty et al. (2014). Shown are the selected covariates of the forward selection procedure for each distributional parameter of the bivariate Student-t distribution. All effects $f(\cdot)$ are modeled semi-parametrically, whereas the remaining effects are included as linear terms.

Table 5

Selected covariates and predictor specification for the bivariate Normal distribution.

Predictor Specification
$\eta^{\mu_{abs}} = Intercept + Growth\_in\_Chinese\_Imports\_1990\_2000 + State\_Income\_Tax\_Progressivity + f(fraction\_of\_Children\_with\_Single\_Mothers) + f(fraction\_Religious) + f(top1percent) + f(fraction\_of\_Adults\_Married) + f(Migration\_Inflow\_Rate) + f(Manufacturing\_Employment\_Share) + f(Social\_Capital\_Index) + f(fraction\_of\_Adults\_Divorced) + f(Household\_Income\_per\_capita) + f(Gini\_Bottom\_99) + f(Teenage\_Labor\_force\_Participation\_Rate) + f(Income\_Segregation) + f(Test\_Score\_Percentile) + f(frac\_with\_Commute\_15\_Mins) + f(Labor\_Force\_Participation\_Rate) + f_{geo}$
$\eta^{\mu_{rel}} = Intercept + State\_EITC\_Exposure + Growth\_in\_Chinese\_Imports\_1990\_2000 + State\_Income\_Tax\_Progressivity + f(Migration\_Outflow\_Rate) + f(Frac\_Foreign\_Born) + f(Test\_Score\_Percentile) + f(top1percent) + f(Fraction\_of\_Adults\_Married) + f(Fraction\_of\_Adults\_Divorced) + f(Income\_Segregation) + f(Fraction\_of\_Children\_with\_Single\_Mothers) + f(Frac\_Black) + f(Social\_Capital\_Index) + f(Household\_Income\_per\_capita) + f(Fraction\_Religious) + f_{geo}$
$\eta^{\sigma^2_{abs}} = Intercept + f(Fraction\_of\_Adults\_Married) + f(Fraction\_of\_Adults\_Divorced) + f(top1percent) + f(Manufacturing\_Employment\_Share) + f(Social\_Capital\_Index) + f(Gini\_Bottom\_99) + f(Household\_Income\_per\_capita) + f(frac\_with\_Commute\_15\_Mins) + f(Racial\_Segregation) + f_{geo}$
$\eta^{\sigma^2_{rel}} = Intercept + f(Manufacturing\_Employment\_Share) + f(Racial\_Segregation) + f(Migration\_Inflow\_Rate) + f(Fraction\_Religious) + f(Income\_Segregation) + f(top1percent) + f_{geo}$
$\eta^{\rho} = Intercept + State\_EITC\_Exposure + f(Teenage\_Labor\_Force\_Participation\_Rate) + f(Household\_Income\_per\_capita) + f(Migration\_Inflow\_Rate) + f(Fraction\_of\_Adults\_Divorced) + f(Frac\_Foreign\_Born) + f(top1percent) + f(Fraction\_of\_Adults\_Married) + f(Racial\_Segregation) + f_{geo}$

Source: Author's calculation based on data from Chetty et al. (2014). Shown are the selected covariates of the forward selection procedure for each distributional parameter of the bivariate Normal distribution. All effects $f(\cdot)$ are modeled semi-parametrically, whereas the remaining effects are included as linear terms.

Table 6
Estimated parametric effects.

	$\eta^{\mu_{abs}}$			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	0.1625	0.0270	0.1097	0.2145
Urban_Areas	0.0450	0.0284	-0.0121	0.1011
Growth_in_Chinese_Imports_1990_2000	-0.0248	0.0086	-0.0421	-0.0078
State_Income_Tax_Progressivity	0.0534	0.0161	0.0230	0.0859

	$\eta^{\mu_{rel}}$			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	-0.0604	0.0292	-0.1149	-0.0015
State_Income_Tax_Progressivity	-0.0050	0.0227	-0.0493	0.0401
Growth_in_Chinese_Imports_1990_2000	0.0107	0.0140	-0.0148	0.0379
State_EITC_Exposure	-0.0213	0.0201	-0.0597	0.0190

	$\eta^{\sigma^2_{abs}}$			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	-0.8800	0.0272	-0.9327	-0.8279

	$\eta^{\sigma^2_{rel}}$			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	-0.4452	0.0364	-0.5151	-0.3774
Urban_Areas	-0.0891	0.0398	-0.1671	-0.0136

	η^{ρ}			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	-0.6874	0.0605	-0.8090	-0.5728
State_EITC_Exposure	-0.0853	0.0389	-0.1576	-0.0056
Urban_Areas	-0.0158	0.0755	-0.1618	0.1335

	$\eta^{n_{df}}$			
	Mean	Std. Dev.	2.5%	97.5%
Intercept	2.4140	0.1070	2.2121	2.6218
Fraction_of_Children_with_Single_Mothers	0.2285	0.1084	0.0292	0.4541
Racial_Segregation	-0.2375	0.1114	-0.4480	-0.0058
Household_Income_per_capita	0.1580	0.1009	-0.0340	0.3578

Source: Author's calculations based on data from Chetty et al. (2014). Shown are posterior mean, posterior standard deviation, as well as pointwise 2.5% and 97.5% pointwise posterior credible intervals of the estimated parametric effects for each distributional parameter of the bivariate Student-t distribution.