

IZA DP No. 9334

Individual Poverty Paths and the Stability of Control-Perception

Hendrik Thiel
Stephan L. Thomsen

September 2015

Individual Poverty Paths and the Stability of Control-Perception

Hendrik Thiel

NIW Hannover and Leibniz Universität Hannover

Stephan L. Thomsen

*NIW Hannover, Leibniz Universität Hannover,
ZEW and IZA*

Discussion Paper No. 9334
September 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Individual Poverty Paths and the Stability of Control-Perception

This paper investigates whether individual control-perception affects the probability of becoming poor, and vice versa, whether poverty experiences can be detrimental to these traits later on. The former relation is intuitive as control related traits underlay many idiosyncratic determinants of poverty. Though traits like control-perception are known to stabilize towards adulthood, the latter association may be plausible when some plasticity is maintained in case of more vigorous environmental influences like poverty. Such deterioration of control-perception would lead to poor people being literally “trapped”. Yet, it is unclear what the underlying mediation paths are and whether control-perception or other potential factors are involved. Our empirical results suggest that poverty experiences affect individual control-perception to some extent. Despite rather modest magnitudes, the findings indicate that no invariance of control-perception is given in adulthood.

JEL Classification: C33, C35, J21, J24, J30

Keywords: personality traits, control-perception, poverty constitution, poverty experience

Corresponding author:

Stephan Thomsen
NIW Hannover
Königstr. 53
D-30175 Hannover
Germany
E-mail: thomsen@niw.de

1 Introduction

Poverty is a global and lasting phenomenon with various manifestations. Whereas poverty is related to appropriate levels of physical subsistence or nutrition for developing nations (see Lambert, 2001), it is defined relative to the incomes of others in developed countries.¹ Albeit a large literature deals with the cross-sectional aggregation and comparison of this concept of poverty (see, e.g., Zheng, 1997), only few studies assess the important dynamic implications of poverty experiences on the individual level (see Aassve et al., 2006, for an outline of the respective literature). This paper therefore examines potential relations between poverty paths and the dynamics of other determinants involved. Apart from control-perception, we assume choice variables like childbearing and household formation, as well as employment to be such interacting entities. Based on longitudinal data from the German Socio-Economic Panel (GSOEP), we use a dynamic structural model that relaxes strict exogeneity assumptions between the model components to some extent. We examine whether control-perception has some direct impact on the development of poverty in terms of different income-based metrics, as well as some indirect effects via other entities involved. Conversely, we also account for a feedback of previous poverty experiences on perceived control and on the mediating variables.

Obtaining a deeper understanding of causal dependencies for individual poverty is insightful for a number of reasons. The main line of argumentation invoked in most debates on poverty builds on the use of occasionally imprecise indicators and often premature inference based on them. This point is best exemplified by annual aggregates of headcount ratios that are the ubiquitous instrument for reportings on poverty (see, e.g., Zheng, 1997), but are also subject to major interpretational and conceptual pitfalls (see, e.g., Foster et al., 2013). Furthermore, headcount aggregates do not take into account how poor the persons concerned

¹ As income is just one means to achieve well-being, a generalized notion of poverty that extends beyond matters of income has been introduced by Sen (1982). It relates to well-being arising from the freedom of choice among potential achievements. However, it is difficult to implement empirically.

are.² Another drawback, especially for the evaluation of causal relations, is that the cross-sectional perspective of poverty aggregates is uninformative in terms of the inter-temporal dimension that poverty evidently possesses and furthermore provides no reasoning on the individual level.³ If poverty were a transitory phenomenon with high turnover rates that bears on different parts of the population over time, a cross-sectional perspective may be adequate. However, as has been shown in various studies (see Stevens, 1999, among others), poor people are often trapped in poverty. Beyond this well documented pattern, the underlying individual causes should be disentangled more explicitly in order to deduce potential counter-measures.

Following the corresponding strands of the literature, two main mechanisms causing such persistence may be in order. First, individuals can differ in terms of characteristics that are relevant for the propensity to experience poverty. Especially when it is assumed that poverty is rooted in income only, the understanding of the relevant causes is well developed and subject to a long-standing literature (see, e.g., Heckman et al., 2006). As of late, the incorporation of cognitive and affective factors stemming from the psychological field (like the control attitudes considered here) adds to this literature (see, e.g., Almlund et al., 2011), also in explaining other outcomes related to labor market success. In economics, such cognitive and affective factors are better known as traits or preferences. Second, affective components and other individual characteristics may be further deteriorated by past poverty experiences, thus locking-in the persons concerned. Such mechanisms have been hypothesized in the sociological literature on poverty for a long time already (see, e.g., Sher, 1977). In empirical economics a reasoning based on changes in attitudes or depreciation in human capital is usually alleged as an implicit explanation for the observed state dependence (see, e.g., Aassve et

² In the literature on axiomatic approaches to poverty, this feature is called distribution-sensitiveness (see, e.g., Zheng, 1997). Further axioms classifying the properties an aggregate poverty measures should comply with are also given in Sen (1976), as well as in Foster and Sen (1997).

³ On the aggregate level, approaches to incorporate dynamic aspects into measures of poverty have been made (see Hojman and Kast, 2009, and the literature they cite). By construction, however, even these dynamic metrics cannot account for individual determinants, as no conditioning sets are accounted for.

al., 2006). As with the above examples, various findings from other disciplines underpin these views and extend the set of potential mediating pathways. A meta-analysis conducted by Haushofer and Fehr (2014) shows that apart from plain economic explanations, like credit constraints, psychological factors (cognitive and affective ones) and even neurobiological factors are evident predictors of poverty traps. For instance, poor living conditions may impede achievements in subsequent tasks via decreased self-regulating capabilities (see Muraven and Baumeister, 2000).

By now, frameworks that allow for a circular dependence between poverty and individual characteristics are bound to a theoretical literature on life cycle saving and wealth accumulation. This particular branch uses concepts from behavioral economics, like hyperbolic discounting, to explain individual heterogeneity in accumulation paths and feedbacks that trap individuals within respective trajectories (see, e.g., Bernheim et al., 2013, for a recent example). Hyperbolic discounting has behavioral implications that can be paraphrased as self-control or self-regulation (see, e.g., Ainslie, 1991). Preference parameters and traits therefore roughly represent the same causes of individual behavior, albeit in different hypothetical frameworks. Preference parameters are utility-related representations of behavioral differences, whereas a trait is seen as more of a task-specific skill or ability in the sense of human capital literature (see, e.g., Almlund et al., 2011). On that account, the various empirical assessment tools that exist in the field of trait psychology (see, e.g., Rotter, 1966, or Tangney et al., 2004) capture different aspects of control-perception, at least to some extent.

To the best of our knowledge, the study at hand is the first one that combines psychometric measures of perceived control and poverty formation in an interacting fashion within a panel framework. It provides an end-to-end treatise along the whole line of argumentation hypothesized by the respective parts of the literature. Using trait measures to explain individual poverty status adds to the literature of poverty constitution, primarily by providing an additional source for typically unobserved individual heterogeneity. In addition, allowing for interdependencies

between both entities further contributes to the literature on general malleability of personality traits throughout adulthood.

The remainder of the paper is organized as follows. The next section gives an overview on definitions and theoretical foundations of poverty in order to establish a suitable notion in the present context. It also works out which mediating factors should be considered within an empirical investigation and which particular role psychological constructs like control-perception may play. The data used in the analysis are introduced in section 3, together with some selected descriptive statistics describing the sample. Section 4 discusses the employed poverty measures and describes the identification of the resulting model specifications as well as the used estimation strategy. The empirical results are presented in section 5. The final section draws a conclusion.

2 Income Poverty and Poverty Dynamics

An initial point to be clarified is as to why an understanding of poverty based on individual valuation or well-being does not always have to coincide with a single-dimensioned lack of income.⁴ For someone to be declared poor or not poor, knowing her or his current income may not be sufficient as the well-being derived from monetary endowments varies across individuals. As such, some preliminary assumptions are required in order to make income a meaningful stand-alone objective. The understanding that underlies a direct association of income and well-being is that income results from rational behavior that seeks to maximize well-being. One possible approach to concatenate both is to use additional data that captures differences in needs, prices, and household composition. Unfortunately, even for individuals that are observationally homogeneous in that regard, preferences, motives and enjoyment abilities are diverse, thus making it still problematic to infer different well-being from individual levels of income. If one allows

⁴ We do not employ the term “utility” in this context as some general utilitarian axioms are unduly strict for the evaluation of income inequality and poverty. Foster and Sen (1997) and much of the related literature elaborate on this criticism. To make this distinction more apparent, we employ alternative terms like “well-being” or “valuation” instead.

for a comparison of individual differences in ratios of well-being however (see Foster and Sen, 1997), it is possible to relate income and well-being via an expenditure function. This setup would require multiple income realizations in a very close time interval (or some stated equivalents, see Dagsvik et al., 2006). Without closeness in time, one runs the risk that constraints and preferences change in between. In most settings (including the present case) such information is not available. As a consequence, it is inevitable to impose some normative assumptions on the individual well-being derived from income.

One possible approach is to impose them on the complete functional form of individual well-being and thus allow for level-comparisons. This understanding of the potential use made from income may be too strict and can be relaxed to some extent. A second possible approach is less narrow and follows from the relativeness of income poverty. In this context, relativeness means that preferences are not claimed to be completely self-interested, but can depend on some distributional reference point. A threshold income that discerns poor and non-poor complies with this requirement. What has to be assumed across individuals is that the difference in well-being that is induced by a certain deficit of the realized income with respect to the reference point monotonically increases as the distance between both income levels grows.⁵ Conversely, a change in well-being arising from a shift towards that reference point has to follow the same rules for all individuals. These assumptions follow the notion of Atkinson's (1970) "ethical observer" in that it is merely assumed that certain hypothetical differences are based on comparable valuations. A rather critical point in this assumption is that the awareness of where this reference point is located also has to coincide across individuals to a very large extent. Otherwise, no judgements about derived well-being can be achieved. Nevertheless, following these presumptions provides a "working definition" that gives individual poverty levels some projection into well-being.

⁵ It should be noted, that the presumptions on the functional form for individual distance-comparisons are somewhat stricter than those originally required by Atkinson's (1970) seminal inequality measure for the aggregate level. This follows from the fact that on the aggregate level, exactly equivalent gains and losses from marginal redistributions of incomes have to be considered, whereas on the individual level with a reference point, varying differences in well-being and varying margins occur at the same time.

2.1 Potential Mediators of Poverty

2.1.1 Socio-Demographic Factors

As already addressed above, the relevance of income differs as the needs of people differ. Many of those needs are objective ones, in that they can be defined by fairly general individual characteristics. Following the family-economic literature, such characteristics may evolve successively or parallel and comprise decisions like household formation, childbearing, and labor market participation (see, e.g., Aassve et al., 2004). They are assumed to take place on an individual basis, but with some collective aims underlying it (see Browning et al., 2011).

The determination of household income is intrinsically rooted in these factors. However, predicting dynamic cross-effects by means of established theoretical frameworks is difficult, as the directions and magnitudes are largely unforeseeable.⁶ For instance, gains arising from household formation may include the ability to exploit economies of scale or comparative advantages in transforming market commodities to household goods (see Becker, 1993). Moreover, the publicness of household goods among household members usually leads to budget increases for further affordable goods. Individuals draw their decisions on these factors to a greater or lesser extent, but owing to their unobserved preferences. Therefore, these features are what the concept of “equivalent incomes” seeks to mimic in empirical investigations of income data. But there are also household characteristics involved in income generation that are not captured by equivalent incomes at all. Several unobserved household patterns may impinge on time constraints or credit constraints, but at the same time may be outcomes of decisions that depend on these constraints. Labor market participation and childbearing are two examples that follow this logic (see Aassve, 2004). The potential to share risk may be another important point in explaining household constitution (see Browning et al., 2011).

⁶ Becker (1993) and Browning et al. (2011) give a comprehensive account on these and related topics.

2.1.2 Preferences and Traits

But also individual characteristics play an important role in terms of poverty risk. Apart from socioeconomic characteristics and other observables that affect incomes and other achievements, certain preferences and traits have similar implications (see Almlund et al., 2011; Thiel and Thomsen, 2013, for overviews). Though the angle of assessment is somewhat different for both concepts, their behavioral dimensions are almost identical. Psychological traits are primarily intended to project various dimensions of behavior into a lower-dimensional continuum, focussing on generality, situation-invariance, and durability. In economics such traits are usually seen as a productivity enhancing human capital stock, where productivity refers to tasks in a wider sense, not only those envisaged on the labor market. Behavioral preference parameters, on the contrary, refer to mathematical laws that link specific stimuli to behavioral responses. In economics, the interest in such parameters is mostly limited to decision and optimization frameworks. An integrative framework for preference parameters and personality traits is yet not explicitly established. Nonetheless, Almlund et al. (2011) provide an overview on some first correlation studies that reveal largely intuitive relationships between both concepts. As such, most of the following findings on the role of preferences and traits in poverty constitution suggest similar mediation paths, though they stem from largely unrelated fields of the economic literature.

An impact of productivity enhancing traits on incomes and related entities is shown in various empirical studies (see, e.g., Heckman et al., 2006). An explicit consideration of poverty constitution, however, is limited to preference-related studies that deal with life-cycle savings. These (mostly theoretical) models attribute interpersonal variation in saving behavior to differences in time preferences, risk tolerance, exposure to uncertainty, and relative tastes for work and leisure, with a particular focus on non-standard types of preferences. They establish that different endowment conditions can lead to individual saving paths that can be understood as a poverty trap. Hyperbolic time preferences as defined by Ainslie (1991), together with borrowing constraints, can lead to occasional exuberance in

consumption, which induces lower wealth-accumulation paths, in turn (see, Laibson, 1997, Bernheim et al., 2013). Such local deviations from individually rational accumulation plans are a form of time inconsistency in preferences, a behavior that Ainslie (1975) has introduced as self-control. Poor people with low assets are more prone to consumption sprees, as the “severity of punishment” is lower for these individuals.⁷

Somewhat related to this notion of executive control or self-control is a person’s so-called “capacity to aspire” (see Appadurai, 2004). In an economic context (see Genicot and Ray, 2012, Dalton et al., 2013), a lack of aspiration can be construed as a factor that endogenously lowers reference points in valuation (relative to agents with higher levels of aspiration) that lead to lower accumulation paths of wealth. There is a circular relation between lower aspirations, wealth levels, and valuations drawn from both. Thus, poverty self-perpetuates in a downward circle, as individuals may lose their aspirations when low income is persistently experienced. There also is a growing empirical support for these mostly model-based mechanisms. Haushofer and Fehr (2014) provide an intriguing argumentation by summarizing experimental and empirical findings from various fields. For one thing, poverty and other unpleasant life events are shown to be causally related to well-being, affect, and stress, where stress levels are gathered through self-information and measured hormone levels. These in turn, are known to have a significant impact on time and risk preferences, building on a substantial literature of behavioral lab-experiments. For another thing, the authors also emphasize that poor people are more liquidity-constrained, making changes in their saving behavior often a matter of external factors rather than of intrinsic preferences and traits. As such, non-normative changes in life circumstances, like a major income drop, impinge on several behavioral parameters, and thus possibly on related traits like control-perception.

⁷ There are some empirical facts underpinning this notion, in that poor people frequently engage in all kinds of commitments in order to stick with their initial saving plans (see, Bernheim et al., 2013, and the literature they cite). For instance, Thaler and Benartzi (2004) show that employee commitments on savings from future wage gains, significantly increased saving rates.

In trait psychology, dynamics over the life course that allow for a comparable reasoning about experiencing major environmental changes are long established, though predominantly for normative ones that are supposed to happen to every person within a certain age span. The corresponding literature shows the highest degree of susceptibility for personality traits in early childhood. From there on, it steadily decreases throughout later childhood, adolescence, and adulthood. For age spans beyond adolescence, large scale cross-sectional analyses show peaks of mean-changes until age 30 (see, e.g., Roberts et al., 2006). These results, however, are moderated when intra-individual measures and very specific or non-normative life events, like death of a spouse, are used. Following Cobb-Clark and Schurer (2013), the effects become even weaker for the working age population, though no complete time-invariance can be established.

Summarizing the above studies, there are several potential associations between income and control-perception, not all of them in a coherent way regarding poor persons and perceived control, though. Some persistent changes seem to have an impact on traits, but are normative in nature and thus also happen to people with higher income. Evidence on non-normative life events, as those that happen to poor or deprived people, are usually based on those events that are onetime occurrences. They may, however, permanently affect the social roles of the people concerned. Thus, the consequences for more persistent but non-normative events, like poverty, are less foreseeable.

3 Data and Descriptives

3.1 The Sample

For our empirical analysis, we use data from the German Social Economic Panel (GSOEP). The GSOEP is a longitudinal survey conducted since 1984 by the German Institute for Economic Research (see Wagner et al., 2007). It provides comprehensive information on a representative sample of German households, including annual information on household income and decision variables related

to household composition and employment. In addition, the survey includes permanent information about labor market history, health, biography, well-being, family background, and living-conditions. In the waves 1994, 1995, 1996, 1999, 2005, and 2010 inventories that measure control-perception are contained as well. As control-perception is among the outcomes of primary interest, the empirical analysis is based on the corresponding waves. The sampling periods in between are used in order to exploit further information on some of the mediating factors. Based on the register of the 2010 wave, a total of 28,776 individual observations are available. As income determination plays a crucial role in analyzing poverty dynamics, we focus on sample members in working age 18 to 65. Considering the timespan from 1994 to 2010, and given panel attrition and unit non-response, we end up with about 13,000 (gross) observations in each wave, where the exact cross-sectional sample sizes feature further drops due to item non-response and varying model specifications.

3.2 Measuring Control-Perception

In order to incorporate perceived control into the empirical model of poverty formation, we employ a trait inventory that comprises questions related to the respective behavioral dimensions. The responses are stated on seven point Likert scales. By use of several items to obtain the individual-specific scores, the reliability of the constructs is generally increased. In the particular case of measuring control-perception, item inventories based on the “Locus of Control” scale of Rotter (1966) are used. The Rotter scale assesses an individual’s attitude on how self-directed (internal) or how coincidental attainments in her or his life are. The underlying trait dimension thus fundamentally relates to the notion of self-control addressed in the previous section, but does not capture exactly the same facets. Self-control is more closely related to concepts from motivational research (like Self-Efficacy). Locus of Control, on the contrary, does only capture beliefs in whether self-determination exists, not in how successful one could be in governing

it.⁸ The GSOEP uses a 10-item version of the original Rotter scale. To make it compliant with an intuitive metric of control-perception, it has to be coded such that high internal (low external) attitudes represent a high degree of control-perception. Identical versions of this scale are available for the waves 1999, 2005, and 2010. A slightly different prequel version was used in waves 1994 to 1996.

Since the trait inventories build on multiple items, relying on unweighted raw scores or arbitrary selections from all available items does not necessarily lead to unidimensional and errorless measures of individual control-perception. To solve the problem of dimensionality, we obtain the finally used item selection from an explorative factor model. To reduce the error proneness, we fit an item response model to the extracted item sets, which in turn is used to obtain latent factor scores for each individual in the sample. Both procedures are applied to all waves that contain control related measures and are detailed and discussed in Appendix B (NOW WEB APPENDIX).

3.3 Descriptive Results: Poverty and Equivalence Incomes

In order to make income a proper indicator of individual well-being it is necessary to adjust income for measurable interpersonal differences in needs. A viable approach in that regard are so-called equivalence weights (see, e.g., Cowell, 2011). We apply a modified OECD scale (see Atkinson et al., 1995). It assigns a weight of 1 to the adult head of a household, a weight of 0.5 to each additional adult member, and a weight of 0.3 to each child being below age 15. The weights are summed for each household in order to obtain the total of equivalent adults that have to share a respective net household income, where household income comprises earned income and capital income. The information in the GSOEP also allows for the consideration of home ownership (i.e. saved rent), social transfers, other transfers, as well annual extra payments. Subsequently, tax payments are computed based on these and other relevant magnitudes.

Using the scale by Atkinson et al. (1995) and designating the cutoff value,

⁸ Control-perception is a major driving force with respect to educational attainments and wages (see, e.g., Heckman et al., 2006, Mueller and Plug, 2006, Heineck and Anger, 2010).

which separates the poor from the non-poor, to be six tenth of the median equivalence income, some first descriptive results are obtained. As illustrated by Figure 1, the poverty line in Germany has increased in nominal terms (on a monthly income-basis) throughout the period from 1995 to 2010. The increase amounts to almost 50 percent, which is only partially on account of an increased price level, as the CPI increase in the same time span is about 25 percent (according to the Federal Statistical Office). Another reason is that some of the skewed frequency mass, especially at the lower tail, shifts to the right between 1995 and 2010, and with it, the reference for the cutoff point.

< Include Figure 1 about here >

Considering the 5-year increments, the increase in the cutoff value has been steady. The corresponding changes in the shares of poor people in the sample are not so, however. They amount to slightly more than 13 percent in 2005 and 2010, and to 9.8 and 9.6 percent in 1995 and 2000. The headcount ratios also do not differ substantially between female and male GSOEP respondents. In case of the 2005 wave, it amounts to 12.7 percent as opposed to 11.1, which is largely in line with ratios provided by census data.

To illustrate the fluctuations among the poor and the non-poor over different timespans, mobility plots as presented in Figure 2 are more suitable. The patterns suggest a substantial degree of persistence for the net equivalence incomes in the different rank-groups over the time intervals 5, 10, and 15 years. Using samples of those who are observed at both points in time, the chances of leaving the lower regions of the distribution seem to slightly increase along with the considered timespans, but the dependence on the initial state is yet tremendous. Even after 15 years, most poor rank-groups are still poor with regard to their equivalence incomes.

< Include Figure 2 about here >

In summary, the descriptive results reveal that path dependence obviously is a major factor, due to reasons whatsoever, and therefore should be analyzed on the

individual level. Moreover, the described ambiguities with regard to the shares of the poor urge for some refined measures in order to better capture the extent of income poverty.

3.4 Descriptive Results: Poverty and Background Characteristics

Selected descriptives on sample characteristics are presented in Table 1 for the 2005 wave. The results for other waves do not differ substantially. In line with the literature on demographic transitions (see Aassve et al., 2004), characteristics that are tied to decision variables underlying household constitution, labor market participation, and the like, are to be considered.

< Include Table 1 about here >

The share of full-time employees is more than twice as large in the male sample. Another substantial divergence holds for full-time job experience and the share of persons that hold a university degree. The means and shares of the remaining variables are relatively equal. Irrespective of their poverty state, about 30 percent of the female and male respondents have a higher secondary schooling degree (not displayed). Roughly 65 percent have completed eight or ten years of schooling. The average age for the sample members is about 40 years. The share of east Germans in the sample largely coincides with the fraction in the overall population. About 68 percent of the respondents live together with at least one child below age 18. About 60 percent live with a partner.

Some mean differences occur when poverty states are considered, however. Only for secondary school degrees, one fails to reject the null of equal mean shares, though only at the 1-percent level for higher secondary schooling in the male sample. For the other characteristics contained in Table 1 some substantial differences between poor and non-poor individuals are apparent, most of them with quite similar patterns for female and male respondents. Most remarkably, the share of full-time employed among the non-poor is more than three times

higher. In line with this, the share of university graduates among the non-poor exceeds that among the poor by almost the same order, but even more for men. Corresponding to the above hypothesis, one finds that poor sample members lack a notable level of control-perception. Females in poverty fall behind by more than 0.4 standard deviations, males even by more than 0.5 standard deviation. For all other characteristics displayed in Table 1, the differences are also sizeable, but to a less extent.

4 Methodology

4.1 Measuring Income Poverty

We derive our poverty metrics on the individual level from different measures on the aggregate level, which build on underlying axioms with well understood implications (see Zheng, 1997). Moreover, much of the usefulness implied by these axioms readily translates to the individual level. Robust inference can only be established if the findings are coherent across different poverty measures. For this purpose, we consider two poverty measures derived from different classes with varying degrees of axiomatic foundation, namely, the headcount ratio, the poverty deficit, and the Watts measure. The selected measures have to comply with the focus axiom (see Zheng, 1997), i.e., they are non-zero only for those individuals who have equivalent incomes below the poverty line L . On an aggregated level, this property has led to the use of right censored income distributions in order to parametrically approximate empirical distributions of poverty. In case of modeling individual magnitudes of poverty, this censoring basically reverses, as measures are zero for non-poor observations and strictly positive otherwise (however, not necessarily continuous).

Let \mathcal{Y}_i be a placeholder for the poverty metrics defined in what follows. Each \mathcal{Y}_i in the sample depends on the corresponding equivalent income y_i and the poverty line L , both assumed to be random variables. Basically, the support for individual equivalence incomes for $i = 1 \dots N$ is the positive real line $\mathbb{R}_{\geq 0}$, but

given that individuals may face different feasible income ranges, the support S^i may vary considerably across individuals.⁹ In terms of the empirical realizations, the support of the union of the N countable collections $S = \bigcup_{i=1}^N S^i$ therefore does not necessarily cover the complete positive real line. The individual poverty metric is a mapping $\mathcal{P}_i(y_i, L) : S^i \times S \rightarrow \mathbb{R}_{\geq 0}$, where the possible realizations of L depend on the exact way in which the mapping is defined.¹⁰ Given that L is determined outside the data generating process that renders the empirical distribution of equivalence incomes F_Y , it may take on any value in $\mathbb{R}_{\geq 0}$. If, however, L directly results from a fraction of a distributional statistic of F_Y (here, six tenth of the median), L is bound to be somewhere in $\{L \in \mathbb{R}_{\geq 0} : L \leq F_Y^{-1}(0.5)\}$.¹¹

For empirical evaluations on the individual level, it is meaningful to preassign exactly one $L = y$ for all N (as we have done in the previous section). The individual magnitude of poverty $\mathcal{P}_i(y_i, L)$ would then change to a conditional measure $\mathcal{P}_i(y_i|L)$. However, the fact that L depends on F_Y , which in turn depends on other $y_j \forall j \neq i$, introduces a problem common to all empirical strategies that model outcomes derived from a distributional statistic of F_Y under *iid* assumption. L is not absolutely independent with respect to the other random variables $\mathcal{P}_j \forall j \neq i$, as all the considered entities are derived from the empirical distribution of y_i .¹² By similar reasoning, each y_i additionally depends on those of potential household members. The necessary change from the joint $\mathcal{P}_i(y_i, L)$ to the conditional $\mathcal{P}_i(y_i|L)$ thus only holds as an approximation. It follows that $\mathcal{P}_i(y_i|L)$ is not exactly *iid*, but gets close to it as N grows. This mild violation of the *iid* assumption has to be tolerated.

The measures $\mathcal{P}_i(y_i|L)$ we use are derived from aggregated poverty measures that are simple (weighted) sums over individual contributions in the sample and thus are easily decomposable. The first one, derives from the headcount ratio (see

⁹ Where the probability measure of each S^i is said to be σ -finite (see Davidson, 1994).

¹⁰ For some cases, e.g. a binary individual contribution as used for the headcount measure, $\mathbb{Q}_{\geq 0}$ (when adjusted for the sample size) or even $\mathbb{N}_{\geq 0}$ would suffice.

¹¹ If L were a quantile and not a fraction of a quantile, it would be restricted to be within the support $S = \bigcup_{i=1}^N S^i$ of F_Y .

¹² To illustrate this point, recall that \mathcal{P}_i can change from zero to some positive value just because another person $j \neq i$ has changed its position in F_Y and thereby affects L .

Sen, 1976) and is defined as

$$H_i(y_i|L) = \mathbb{1}(y_i \leq L). \quad (1)$$

The second one derives from the poverty deficit (see Lambert, 2001) and has the virtue to account for the magnitude of poverty as well. It reads

$$PD_i(y_i|L) = (L - y_i)\mathbb{1}(y_i \leq L). \quad (2)$$

The third alternative is, when aggregated over observations, the only measure considered here which is completely distribution sensitive. It has been established by Watts (1968) and is closely related to the entropy concept employed in information theory (see Theil, 1967). The individual-specific contribution reads as follows

$$W_i(y_i|L) = (\log L - \log y_i)\mathbb{1}(y_i \leq L).^{13} \quad (3)$$

Apart from the decomposability, the latter two measures also quantifies the distance that was established as a necessity for an interpretation in terms of well-being in Section 2. In addition to the mentioned focus axiom, the headcount and the Watts measure also shares the property of scale invariance (see Zheng, 1992). The headcount ratio is also characterized by location invariance, a property that no distribution sensitive poverty measure fulfills in general (see Zheng, 1994). Scale invariance implies that a common factor applied to the y_i of all poor individuals, does not change the aggregate measure. It translates into the individual specific contributions as well. However, complying with scale invariance does usually not suffice to account for price level changes over time, except when exactly the same share of income is affected by the price level change for all poor individuals. As even the most basic commodity bundles represent different relative shares of the respective overall incomes, this assumption is unreasonable though. As such, price level changes should be considered for the computations of the poverty measures on the individual level.

¹³ In its aggregated form, the Theil entropy measure for all N with $y_i \leq L$, $T(y|L)$, enters the Watts measure by $W(y|L) = H(y|L) \left[T(y|L) - \log\left(1 - \frac{PD(y|L)}{H(y|L)L}\right) \right]$.

4.2 Identification and Consistency

Keeping dependencies on y_i implicit, consider $D(\mathcal{Y}_{i|L})$ to be a parametric distribution that properly represents the individual contribution to one of the respective poverty measures addressed in the previous section, where $\mathcal{Y}_{i|L}$ is a placeholder for the measure-specific scalar random variable.¹⁴ For instance, in case of the contribution to the headcount measure, the distribution for $\mathcal{Y}_{i|L}$ would be Bernoulli with respective conditional expectation and link function (see McCullagh and Nelder, 1989). As the considered mediating pathways suggest, it is important to account for three features that impinge on the model structure in a dynamic perspective: (i) the path/state dependence of individual poverty formation, (ii) potential feedbacks from the current poverty states to at least some determinants of poverty in the future, (iii) the initial conditions of the poverty paths at the beginning of the sampling period.

4.2.1 State Dependence and Lagged Feedback

In order to properly account for a poverty-trap, some kind of state dependence for the poverty measure under study has to be introduced into the empirical model. For this purpose, a first order autoregressive process for the outcome variable is sufficient, as the interest is not in a complete representation of the individual paths over a large timespan. Moreover, it has to be considered that at least some individual determinants underlying poverty are not independent of previous poverty experiences, as it is likely that past poverty experiences further depreciate those individual characteristics. Such behavior, which Wooldridge (2000) terms a feedback, implies that the development of some explanatory variables $Z = (z'_1, \dots, z'_T)$ can be considered to take place outside the model throughout the whole sampling period, whereas for variables that are subject to feedback this only holds for some sampling periods. For every period t , the latter are contained in the vector w_t . Moreover, the panel structure of the data allows for the incorporation of some otherwise unobserved individual heterogeneity c_i that is assumed to be time invariant.

¹⁴ All of the identification results extend to more general parameterizations of $D(\cdot|\cdot)$, i.e., to other poverty measures not considered here.

Given this distinction, the respective distribution of individual poverty measures $\mathcal{Y}_{it|L}$ conditional on covariates z_{it} and w_{it} , as well as on unobserved heterogeneity c_i , reads

$$D_t(\mathcal{Y}_{it|L}|w_{it}, z_{it}, x_{it-1}, c_i), \text{ with } t = 1, 2, \dots, T \text{ and } x_{it} = (\mathcal{Y}_{it|L}, w'_{it}).^{15} \quad (4)$$

Treating c_i as an incidental parameter to be estimated causes severe consistency problems (see Neyman and Scott, 1948). Giving an explicit account on c_i has some clear advantages over this. Following the approaches of Mundlak (1978) and Chamberlain (1982), one can parameterize c_i conditional on covariates. Modeling c_i in that way eludes arbitrary dependence among the error terms of $\mathcal{Y}_{it|L}$ and does not restrict observed and unobserved factors to be independent, i.e., $w_{it}, z_{it} \not\perp c_i$ is allowed for. However, the lagged dependent part of x_{it-1} in equation 4 depends on c_i by construction. Putting aside this dependence for the moment, one can formally restate the above arguments on Z_i as a requirement that each z_{it} is strictly exogenous, implying that $D_t(\mathcal{Y}_{it|L}|z_{iT}, z_{iT-1}, \dots, z_{i1}, c_i) = D_t(\mathcal{Y}_{it|L}|z_{it}, c_i)$.¹⁶

Recall that the vector w_{it} ($\forall t = 1, 2, \dots, T$) contains the mediating characteristics of poverty along with perceived control. Much like z_{it} , the elements of w_{it} are driving forces of poverty, but they are deemed to be affected by past poverty states. Besides perceived control, outcomes like childbearing, household formation, and employment are assumed to be affected by a similar reversion. We have initially stated that such feedbacks urge a partial relaxation of the strict exogeneity assumption. In the terminology of Engle et al. (1983), w_{it} is predetermined with respect to $\mathcal{Y}_{it|L}$ for $t - 1, \dots, 0$, implying that for each t , w_{it} is independent of the current and future error terms $s \geq t$ of $\mathcal{Y}_{it|L}$.

This relaxation complicates the modeling of the joint distribution $\prod_{t=1}^T D_t(\mathcal{Y}_{it|L}|$

¹⁵ Note that without additional requirements, higher order lags of $\mathcal{Y}_{it|L}$ and w_t could be included. Then the conditional distribution changes to $D_t(\mathcal{Y}_{it|L}|w_{it}, z_{it}, X_{it-1}, c_i)$, with $X_{it-1} = (x_{it-1}, \dots, x_{i1})$ and $x_{it} = (\mathcal{Y}_{it|L}, w'_{it})$.

¹⁶ In terms of conditional expectations $E(\mathcal{Y}_{it|L}|z_{iT}, z_{iT-1}, \dots, z_{i1}, c_i) = E(\mathcal{Y}_{it|L}|z_{it}, c_i)$. According to the definition of Engle et al. (1983), z_{it} is also weakly exogenous such that its data generating process takes place outside of the conditional model in equation 4, without any overlap in the parameter vectors. It is thus possible to refrain from any further discussion on the marginal distributions of $z_t \forall t$.

$w_{it}, z_{it}, x_{it-1}, c_i$), as one cannot apply the same simplification as in case of Z_i , or z_{it} respectively. Without w_{it} , it would suffice to properly account for the initial poverty state in $t = 0$ to make the joint distribution a product of the T conditionally independent distributions $D_t(\mathcal{Y}_{it|L}|w_{it}, z_{it}, c_i)$. In presence of the feedback effect on w_{it} , this property no longer holds (see, e.g., Arellano and Honoré, 2001). Given the set of properties discussed for the time paths of $\mathcal{Y}_{it|L}$, z_{it} , and w_{it} thus far, two frameworks that can consistently estimate the parameters of interest may be considered.

Partial Likelihood Approach: One possible solution is to refrain from any independence assumption discussed within the last paragraph, and thus from any assumption on the joint distribution of the individual paths over T . Instead, one merely has to settle for the correct specification of the period-specific distributions $D_t(\mathcal{Y}_{it|L}|\cdot)$ for all $t = 1, \dots, T$. If these period specific distributions are correctly specified and treated like distributional contributions in a pooled sampling context, strict exogeneity is not a necessary condition for consistency any longer. This finding builds on a special case of general consistency results in presence of partial misspecification for maximum likelihood and extremum estimators (see White, 1982). Following the Kullback-Leibler identity, it can be shown that averages over single factors of a joint distribution suffice in order to establish consistent estimates. In the case of averaging over the joint distribution along the time dimension, Wooldridge (2002) calls this a partial likelihood approach. However, it cannot jointly quantify the dynamic interactions between $\mathcal{Y}_{it|L}$ and w_{it} , as would be the case given more structure along the time dimension. Moreover, it should be noted that contemporaneous exclusion restrictions among some of the possible combinations of the variables in w_{it} have to be imposed. As opposed to the case where the equations for $\mathcal{Y}_{it|L}$ and w_{it} are to be considered simultaneously, unrestricted contemporaneous cross-effects are not a matter of identification. Instead, they would lead to some kind of self-imposed simultaneity bias. One thus still has to make sensible choices about which elements of w_{it} contemporaneously enter the partial likelihood models for other elements of w_{it} . As the order cannot be

empirically inferred, one has to base the restrictions on economic theory.

Structural Approach: Another empirical approach pursued in the present setting is a structural one that relaxes the strict exogeneity assumptions for w_{it} . It jointly models the entities contained in $\mathcal{Y}_{it|L}$ and w_{it} . It builds on the findings discussed in Wooldridge (2000), who suggests to factorizes the individual processes for $\mathcal{Y}_{it|L}$ and the set of predetermined covariates w_{it} , $x_{it} = (\mathcal{Y}_{it|L}, w'_{it})$. If one assumes that, in addition to strict exogeneity with respect to $\mathcal{Y}_{it|L}$, z_{it} is also strictly exogenous with regard to w_{it} , one can write

$$D(x_{it}, \dots, x_{i1} | z_{iT}, \dots, z_{i1}, c_i) = \prod_{t=1}^T D_t(x_{it} | z_{it}, x_{it-1}, c_i) \text{ with factorization} \quad (5)$$

$$D_t(x_{it} | z_{it}, x_{it-1}, c_i) = D_t(\mathcal{Y}_{it} | w_{it}, z_{it}, x_{it-1}, c_i) D_t(w_{it} | z_{it}, x_{it-1}, c_i).$$

Assuming that all conditioning variables in equation 5 enter the distributions of $\mathcal{Y}_{it|L}$ and w_{it} in a linear-additive fashion and given a corresponding link function, standard identification theory based on cross-equation restrictions, exclusion restrictions, and covariance restrictions can be applied in order to render the model identified. However, given the particular mixture of linear, binary, and corner solution link functions that arise from the variable-types in $\mathcal{Y}_{it|L}$ and w_{it} , some peculiarities compared to the linear case are in order. These requirement kind of predesignate the first identification restriction. As shown by Maddalla (1983), all systems of binary or censored endogenous variables (or mixtures of them) should be recursive with respect to contemporaneous cross-effects. Omission of this recursive design leads to the case where at least some of the equations involved are logically inconsistent, i.e., the sum over all joint probabilities do not generally sum to one. Recursiveness implies logical consistency, but is not a necessary condition in all possible realizations.¹⁷ If we impose no restrictions on the equations

¹⁷ For corner solution equations like the poverty deficit and the Watts measure, logical consistency depends on specific parameter realization and restrictions may be weaker than recursiveness. The necessary and sufficient conditions on the parameter space of the contemporaneous endogenous variables would not be feasible as a reparameterization, but only as an inequality-constraint optimization. This is relatively impractical and, furthermore, the resulting model has no meaningful economic interpretation. For binary link functions involved, however, the recursiveness assumption is strictly necessary.

for $\mathcal{Y}_{it|L}$, the recursiveness assumption in the adjacent equation in w_{it} is mathematically equivalent to the requirement for predeterminedness of this mediating variables with respect to $\mathcal{Y}_{it|L}$. Off course, for logical consistency, recursiveness and thus predeterminedness have to extend to the contemporaneous cross-relations among all further variables in w_{it} as well. It follows that the contemporaneous cross-effects have to decrease row-wise.

For complete identification of the simultaneous structure in equation 5, we have to introduce a second type of restriction. Since we explicitly model the unobserved effects, we opt for cross-equation covariance restrictions among the residuals. As c_i is properly accounted for and is allowed to vary by equation, it does not seem too restrictive to do so. Alternatively, exclusion restrictions on the respective z_{it} -vectors could be imposed, but justifying the required instrument is a more difficult task in the current setting.

4.2.2 Initial Conditions

Irrespective of using the partial likelihood or the structural approach to allow for predeterminedness, the initial poverty status for the start of the sampling period in $t = 0$ has to be addressed. For dynamic panel data models with rather small T , misspecified initial conditions $\mathcal{Y}_{i0|L}$ and w_{i0} are serious confounders for parameter consistency, as opposed to time series frameworks with large T . Treating the initial conditions as a non-stochastic component would also imply that they are not allowed to depend on heterogeneity c_i , which is not very plausible. If the initial conditions are assumed to be stochastic, Hsiao (2003) discusses cases of equilibrium initial conditions that allow to retrieve their distribution functions and to consider them as part of the joint distribution in equation 5, rather than as a conditioning variable. However, such presumptions are not testable in practice and it is unlikely that the starts of the processes $\mathcal{Y}_{it|L}$ and w_{it} always coincide with the start of the sampling period. We use an approach introduced by Wooldridge (2005), instead. It models c_i as a function of $\mathcal{Y}_{i0|L}$, the elements of w_{i0} , the individual specific time

averages \bar{z}_i , and a remainder of unobserved heterogeneity a_i , implying

$$D(c_i | \mathcal{Y}_{i0|L}, w_{i0}, \bar{z}_i, a_i), \tag{6}$$

where the components $\mathcal{Y}_{i0|L}$, w_{i0} , \bar{z}_i , and a_i are linear and additive. Given this specification, the initial conditions are not part of the joint distribution. Instead, by solely conditioning on $\mathcal{Y}_{i0|L}$ and w_{i0} , one can remain unconcerned about the distributions of the initial conditions. The distribution $D(\cdot|\cdot)$ is chosen to coincide with that of the respective outcome $\mathcal{Y}_{i|L}$ or w_i , where for normal-based distribution types both terms conflate to one linear-additive condition set.

4.2.3 Sample Spacing

One additional problem in the current setting is imposed by the fact that control-perception is not sampled in even intervals. Without formal derivation, it is immediately obvious that the models considered thus far cannot consistently estimate the state dependencies within the paths of poverty experiences and predetermined variables when sampling periods t are unequally spaced.¹⁸ That being the case, the reference period for the underlying data generating process, usually termed the unit period (see Fuleky, 2012), does not coincide with the observational interval. Approaches that account for these issues (see Baltagi and Song, 2006, for an overview) are not applicable to non-linear dynamic settings. As such, we treat the problem by setting up different subsets of the data with varying but equally spaced sampling gaps and cross-validate the results derived from them.¹⁹

¹⁸ A formal representation is given in Millimet and McDonough (2013).

¹⁹ It should be noted that equal observational intervals also represent an irregular spacing regarding the unit period and the data generating process. This follows from the fact that the unit period at which the individual is supposed to make consecutive decisions almost never complies with the rate at which the sampling occurs (e.g., annually). It can be shown that the state dependence parameter of the true process mixes with the error term of the observed model in this case (see Millimet and McDonough, 2013). The resulting estimates are consistent, but actually with respect to the “wrong” model parameters. Given equal spacing, the misspecification can be regarded as being constant, though. This still allows for meaningful inference.

4.3 Parameter Estimation

For the structural approach, the aforementioned focus on the labor force, i.e., on individuals aged 18 to 65, implies to retain only those individuals in the sample that are in working age for the complete time path to be considered. Given time paths of length T years including the initial period, all observations in the initial period are aged between 18 and $65 - T$, whereas in the last observational period the age varies between $18 + T$ and 65. On the one hand, this proceeding has the virtue of decreasing the relative weight of probably altered transition periods out of the labor market, since only the last sample waves get close to the legal retirement age. On the other hand, rather practical contemplations underly this step: The structural approach requires contiguous individual time paths to set up the likelihood contribution and only few waves provide information on perceived control. The partial likelihood approach is less “data hungry” as only two adjacent intra-individual observations are needed in order to obtain a consistent partial likelihood contribution. Thus, the number of observations is generally higher for the pooled models.

Moreover, we consider gender specific subsamples for our analysis. This has the intuitive reasoning that human capital pricing and thus income, as well as labor market participation and other factors, differ by gender. It also greatly simplifies the underlying structures for the estimation and the computation of the standard errors, since there is relatively little need to account for intra-household correlations. The samples for male and female respondents are quite distinct in that regard.²⁰

²⁰ Almost 89 percent of both gross samples do not live together with another sample member who is in working age and of the same gender in 2010. If only those observations without missing values in the variables of interest are retained, this share increases to above 99 percent in either case. As such, the dependence structures within the individual time paths seem to be the only ones of actual importance. The gender subscripts are kept implicit in the following formal representations.

4.3.1 Partial Likelihood Approach

Though not being jointly estimated, the predeterminedness among the variables in w_{it} follows the same order as derived for the structural model. Likewise, the link functions are identical to those in Table 2. The estimation of the partial likelihood models is fairly easy to implement as an equation-by-equation pooled estimator.

Recall the vector $x_{it} = (\mathcal{Y}_{it|L}, w'_{it})$ combining the respective poverty measure with the predetermined mediating factors and perceived control from equation 4. The partial likelihood approach discussed in the previous section separately estimates the respective equations for all K variables in x_i^k ($k = 1, \dots, K$). Each variable x_i^k can be associated with a respective link function that characterizes its conditional expectation, and hence, its probability distribution. The link functions corresponding to the variables x_i^k are summarized in Table 2.

< Include Table 2 about here >

The dependent variables are conditioned on lagged values x_{it-1} , on strictly exogenous variables z_{it} , and on the unobserved heterogeneity term $c_{i|\mathcal{Y}_{i0|L}, \bar{z}_i}$, or $c_{i|w_{i0}, \bar{z}_i}$ respectively.²¹ As stated above, contemporaneous cross-effects among the elements of x_{it} cannot be arbitrarily specified, as the estimates are otherwise inconsistent due to a self-defined simultaneity. Given the hypothesis that the feedback effects disseminate from past poverty to perceived control with all other elements of w_{it} being mediating factors, it is self-evident to allow $\mathcal{Y}_{it|L}$ to be contemporaneously affected by all w_{it} . By the same token, perceived control is the K th element of w_{it} with no contemporaneous cross-effects. For the remaining variables in w_{it} , the order of the contemporaneous cross-effects are ad hoc choices that cannot be based on the data at hand. Instead, economic theory suggests that household formation with a partner usually takes place before childbearing decisions are made. We follow this convention here. The positioning of employment is more complex from a theoretical perspective. For women, childbearing is known to negatively affect labor force participation and thus employment (see Aassve et al., 2006). For men,

²¹ Note that in case of the partial likelihood approach, the explicit consideration of a time invariant remainder term a_i is meaningless as no time paths are modeled. Thus, a_i can be absorbed into the time-specific error term.

on the other hand, labor market participation and employment may be more of an preliminary decision, as employment is a promoting factor in mating and search frameworks (see Burdett and Coles, 1999, Aassve et al., 2002). We will test these presumptions under the structural approach.

To give a more ostensive representation of the partial likelihood specification, consider the case of the binary headcount $\mathcal{Y}_{it|L} = H_{it}$ as a left-hand side example for x_{it}^1 . Then, the explicit representation of $D_t(\cdot|\cdot)$ is

$$\Phi [(2H_{it} - 1)(\beta'_1 z_{it} + \beta'_2 \tilde{w}_{it} + \beta_3 H_{it-1} + \beta'_4 \tilde{w}_{it-1} + \alpha_1 H_{i0} + \alpha'_2 \bar{z}_i)],$$

with z_{it} and \bar{z}_i having ones as their respective uppermost element. The implementation for the other outcome equations follows the same logic. The resulting log-likelihood contribution for each x_i^k -specific pooled model is

$$\ell_i(\Gamma_k) = \sum_{t=1}^T \ln D_t(x_{it}^k | z_{it}, \bar{z}_i, \tilde{w}_{it}, x_{i0}^k, \Gamma_k),$$

where \tilde{w}_{it} is always a $(K - 1)$ -subset of w_{it} , except for $x_i^k = \mathcal{Y}_{i|L}$, due to the otherwise arising simultaneity problems. Again note that the partial likelihood approach does not explicitly involve the unobserved component a_i . Instead, it is absorbed into the respective error term. This affects the scale normalization for binary models or the variance estimate in the censored and linear case. In all three cases, however, the implied serial error-correlation on the individual level has to be accounted for when standard errors are to be computed.

4.3.2 Structural Approach

As has been argued to establish the identification of the model, the use of time-invariant random effects makes the assumption of zero covariances across equations plausible. By the same reasoning, the individual-specific joint distribution over time can be assumed to require no further free form correlation in the idiosyncratic error terms. Without such correlations, the likelihood derived for the estimation of the structural model can be evaluated without any multidimensional integrals.

This presumption is not necessary for identification, but greatly alleviates the estimation procedure. Following this condition and given the identification results established in the previous section, we can write the joint distribution of $\mathcal{Y}_{it|L}$ and the K row elements of w_{it} over the sampling period as a simple product

$$\begin{aligned}
& D(\mathcal{Y}_{i1|L}, \dots, \mathcal{Y}_{iT|L}, w_{i1}, \dots, w_{iT} | z_{i1}, \dots, z_{iT}, \mathcal{Y}_{i0|L}, w_{i0}, c_i, \Gamma) \\
&= \prod_{t=1}^T D_t(\mathcal{Y}_{it|L} | w_{it}, \mathcal{Y}_{it-1|L}, w_{it-1}, z_{it}, \mathcal{Y}_{i0|L}, w_{i0}, c_i, \Gamma_1) \\
&\quad \vdots \\
&\cdot D_t(w_{Kit} | \mathcal{Y}_{it-1|L}, w_{it-1}, z_{it}, \mathcal{Y}_{i0|L}, w_{i0}, c_i, \Gamma_K),
\end{aligned} \tag{7}$$

where the partitions Γ_k are generally not the same as in case of the partial likelihood approach above. In order to maintain a comparably sparse parameterization, we do not allow all the parameters in $D(c_i | \mathcal{Y}_{i0|L}, w_{i0}, \bar{z}_i, a_i)$ to vary across equations, but use an overall scaling factor for c_i in each equation. For the first equation that generally models $\mathcal{Y}_{it|L}$, the scaling factor is always one. As such, the parameter blocks in $\Gamma = (\Gamma'_1, \Gamma'_2, \dots, \Gamma'_K)$ have the parameters for c_i in common. The link functions for the respective $\mathcal{Y}_{it|L}$ and w_{it} are analogous to those summarized in Table 2.

For a better illustration of the specifications resulting from equation 7, consider again the case of the binary headcount $\mathcal{Y}_{it|L} = H_{it}$, for simplicity only along with perceived control as a scalar predetermined variable $w_{it} = \theta_{it}$. Then one obtains individual time paths

$$\begin{aligned}
& \prod_{t=1}^T \Phi [(2H_{it} - 1)(\beta'_1 z_{it} + \beta_2 \theta_{it} + \beta_3 H_{it-1} + \beta_4 \theta_{it-1} + \psi + \alpha_1 H_{i0} + \alpha_2 \theta_{i0} + \alpha'_3 \bar{z}_i + a_i)] \\
& \frac{1}{\sigma} \phi [(\theta_{it} - \delta'_1 z_{it} - \delta_2 H_{it-1} - \delta_3 \theta_{it-1} - \delta_4 (\psi + \alpha_1 H_{i0} + \alpha_2 \theta_{i0} + \alpha'_3 \bar{z}_i + a_i))(1/\sigma)].
\end{aligned}$$

It is implied by our hypothesis that control-perception is always the lowermost equation in the system 7, i.e., it is the variable that is always predetermined with respect to all other dependent variables at each t . Likewise, $\mathcal{Y}_{i1|L}$ is always the variable that is allowed to be contemporaneously affected by all w_{it} , and thus is

always the uppermost equation in the system. The remaining endogenous variables in w_{it} may follow an order of predeterminedness established by the same economic reasoning as in the case of the partial likelihood approach discussed above. The simultaneous estimation pursued here provides the opportunity to nest statistical testing procedures in order to extract such information from the data.²² We use a general specification test for simultaneous equation systems suggested by Anderson and Kunitomo (1992). It tests for predeterminedness against the alternative of unrestricted cross-effects among the elements of $\mathcal{Y}_{i1|L}$ and w_{it} . This choice is rooted in one particular limitation imposed by the setting at hand. Following the identification and consistency considerations addressed above, predeterminedness has to be imposed for logical consistency. As such, it is only possible to derive test statistics from (sub-)models under this assumption, since the unrestricted model is logically inconsistent given the above arguments.

Having solved the issues of predeterminedness and logical consistency, what remains to be addressed is how to treat the time invariant unobserved component a_i . By assuming that $a_i \sim \mathcal{N}(0, \sigma_{a_i})$, the following log-likelihood contribution for individual i over the sampling periods T is obtained

$$\ell_i(\Gamma_1, \dots, \Gamma_K) = \ln \int D(\cdot | z_{i1}, \dots, z_{iT}, \bar{z}, \mathcal{Y}_{i0|L}, w_{i0}, \dots, w_{iT}, a_i, \Gamma_1, \dots, \Gamma_K) \cdot C da,$$

where $C = \left(\frac{1}{\sigma_a}\right) \phi\left(\frac{a}{\sigma_a}\right)$ and $D(\cdot|\cdot)$ is the right-hand side product of equation 7. The integral over the unobserved a_i can be solved numerically by means of a Gauss-Hermite quadrature. The number of interpolation nodes required for obtaining a relatively accurate approximation result is relatively low in cases where $D(\cdot|\cdot)$ involves a link function based on a normal distribution (see Butler and Moffitt, 1982), which applies to all the link functions in Table 2.

²² Unfortunately, such procedures are rare and largely limit to time series applications with large T (see, e.g., Kilian and Vega, 2011).

5 Results

The main results presented in the following section are based on the 5-year sampling interval. The 1995 wave represents the initial period, whereas the waves 1999 to 2010 model the actual individual specific time paths. Other observational intervals and specifications are presented in Section 5.3. For all results to be considered, note that most of the average partial effects possess a self-explaining magnitude. For perceived control, all effect sizes refer to a change on its standard deviation or from its standard deviation. For the poverty deficit, the average effects can be interpreted in terms of the absolute distance of equivalence incomes to the poverty line. Similarly, for the Watts measure the average change refers to the logs of both entities. Hence, one should always put into perspective that average partial effects for the poverty deficit measures tend to be rather large in magnitude, as they relate to induced changes in equivalent euros. On the other hand, model parts that include the poverty deficit as a right-hand side variable produce comparably small average effects as opposed to those with binary indicators of poverty involved, though their absolute meaning may be quite substantial. The effects for the strictly exogenous covariates are not discussed in the following due to space considerations. For the Watts measure a meaningful interpretation is somewhat more difficult to establish. If the deficit in logs is a left-hand side variable, one may reformulate the average partial effects of continuous variables by means of the exponential function. The resulting average partial effects then refer to the implied average change on the ratio of the poverty line and the net equivalent income. By the same transformation, the average effect sizes of the Watts measure as a right-hand side variable are in terms of a one percent increase in the ratio of poverty line and equivalence income with respect to the corresponding dependent variable. For discrete explanatory variables, no direct transformation for the log ratio is available, as it also depends on its level when the change is discrete instead of marginal.

5.1 Partial Likelihood Approach

The empirical results for the female sample using the partial likelihood approach are presented in Tables 3 to 5. For males, the corresponding results will be presented in Tables 6 to 8. At the first glance, it becomes obvious, that state dependence plays a dominant role in the model parts for poverty status, full-time employment, living with partner, and childbearing.²³ Being poor in the previous period vastly increases the probability of living in poverty in the ensuing one. The same holds for full-time employment status and the other considered predetermined variables. Being currently full-time employed also significantly reduces the probability of contemporaneously living in poverty.

< Include Table 3 about here >

Regarding the exact magnitudes of the presented estimates, one finds that, probably owing to the large sample size, a lot of statistically significant effects are at hand. The effect sizes for the strictly exogenous variables are largely in line with what could have been expected based on economic rationales. In addition to the contemporaneous exogenous effects, the significant coefficients of the time-averaged indicators suggest the prevalence of some characteristics that are highly correlated with average (unobserved) behavioral driving forces that go beyond perceived control, such as intelligence, further unobserved abilities, and motivational factors. Adding up the time-invariant and time-varying components of the strictly exogenous variables, one finds that having some secondary school degree reduces the average probability of living in poverty by roughly 7 percentage points, obtaining a higher secondary degree does so by even 8 percentage points. Similarly sizeable is the 6 percentage point reduction in probability when holding a university degree. Discarding the impact of those time-invariant variables that act as an indicator for unobserved heterogeneity, having some vocational qualification and holding German citizenship also significantly contribute to the explanation of individual poverty states. Those individuals who possess a vocational degree

²³ We will use “childbearing” as a synonym for “having at least one child” throughout the following discussion.

are almost 4 percentage points less likely to live in poverty than those who do not. Holding the German citizenship lowers the probability of living in poverty by roughly the same magnitude, whereas living in the eastern part of Germany has an opposing effect that is twice that size. Moreover, though being jointly significant, the coefficients of age included as contemporaneous regressors in the poverty equation do not show any clear pattern.

Considering the contemporaneous dependent variables, employment status and living with a partner decrease the probability of the binary poverty status by 6.9 and 6.5 percentage points, respectively. This seems quite intuitive. The effects of the lagged characteristics are not completely in line with what could have been expected. On the one hand, having been gainfully employed in the previous period significantly reduces the poverty risk in a given period. On the other hand, living together with a partner at $t - 1$ increases the risk of being poor, though by a comparably small margin of one percentage point. The same holds for having had at least one child in the previous period. The likely explanation for these somewhat contradictory effects might be that, after controlling for potential economies of scale by using equivalence incomes, working-age individuals who live together with others may not solely benefit from living with each other. Quite often, such individuals may be sole earners or at least have to keep one additional household member. This finding is somewhat at odds with those results derived from traditional random effects models under strict exogeneity assumptions (see, e.g., Biewen, 2004), implying that its relaxation is a quite reasonable step in the setting at hand.

The estimate for the state dependence effect is the strongest one. Having been poor at $t - 1$ raises the probability of being poor in the subsequent observational period by about 19 percentage points. This effect highlights that even after controlling for differences in observed and unobserved characteristics, past poverty experience is connected to a higher future poverty risk. The revealed state dependence corresponds to the previous empirical findings that have been discussed throughout the review of the related literature. The fact that the incorporation of

perceived control into the model does not change this pattern indicates that this particular trait does not add much to the explanation of this implicit association. Furthermore, it should be noted that given the current setting those who are poor at $t - 1$ and again at t may consist of two rather different groups. There are those individuals for whom the two points in time are part of a continuing poverty spell. Additionally, there may be those individuals who have an interrupted spell of poverty, or potentially even more than one. In the setting at hand, a potential mixing of these groups is even more likely as the observational points in time are quite distant. Given that the partial likelihood approach does not distinguish between likelihood contributions across individuals and within individuals along the time axis, this issue is not properly accounted for by the results presented here. The implications of continuing spells and repeated poverty unemployment may be somewhat different. One may learn more about this phenomenon from the data when the observational interval as well as the modeling of the individual time paths are changed. This will be subject to Section 5.2 and Section 5.3.

The estimation results presented thus far remain valid when the two remaining poverty measures, namely the poverty deficit and the Watts measure, are considered. Recall that, as opposed to the binary indicator, both measures also capture the extent of poverty, where the Watts measure puts more weight on equivalent incomes in further distance to the poverty line. The corresponding estimation results are presented in Tables 4 and 5. The order of effect sizes presented for the binary poverty indicator thus far do not change for either measure. Bear in mind that the average effects refer to the conditional expectation for the complete sample, not only to those observations for whom the deficit measures are not censored. Being employed reduces the average deficit by roughly 200 equivalent euros, or given the Watts specification, decreases the log-ratio between the poverty line and the equivalent income by 0.4 percentage points. Analogously to the binary case, living with a partner and the degree of perceived control also exert a substantial effect in terms of poverty reduction. For the strictly exogenous variables, the picture is slightly different compared to the binary case. Educational achievements, which

have been a strong predictor for the headcount measure, are less significant for the poverty deficit and the Watts measure. Holding a vocational degree, on the contrary, seems to decrease the poverty deficit by 60 equivalent euros on average, or in terms of the log ratios by 0.12 percentage points. Apart from a rather small impact of age, the time-averaged covariates in the correlated part of the model have lost most of their statistical significance given the two specifications that involve the poverty deficit and Watts measure.

< Include Tables 4 and 5 about here >

Quite naturally, the results for the predetermined left-hand side variables do not depend on the specific poverty measure being employed.²⁴ Some of them are somewhat remarkable, though. Having at least one child negatively affects the probability of employment, as does living with a partner. Perceived control has some positive effect on employment, but the magnitude is negligible. Again, the state dependence in the respective dependent variable is the most influential predictor. Having been employed in the previous observational period raises the probability of employment in the current period by 31 percentage points. Similar findings occur in case of having one or more children, and in case of living with a partner. These amount to 50 and 21 percentage points, respectively. It should be noted, however, that as with the state dependence in poverty, two rather different groups of observations are likely to mix up in the constitution of the state dependence in the employment equation. Again, there may be those observations with the two points in time being part of a continuing spell without employment, and those who have one or more intervening spell(s) of employment. As argued above, a model that considers complete individual time paths may contribute to a better understanding in this case. With regard to the exogenous variables, having a university degree is a particularly strong exogenous predictor for employment. The same holds for possessing a degree from a technical college. Considering the time-invariant part of the model, these effects are mitigated to some extent. For

²⁴ The possible changes in the estimated coefficients, if anything, affect the rightmost reported decimal digit.

all further strictly exogenous variables, the effects are rather moderate in magnitude across all columns. The results for perceived control are quite plausible and in large parts in line with the respective literature on stability. As already stated, they exhibit the anticipated effects on current employment status. The effects are positive in that a higher degree of perceived control increases the probability for those outcomes typically associated with labor market success (see Almlund et al., 2011). As such, it is also possible that an additional impact on poverty status transmits via the effect on employment status. The estimates for the lagged effects of perceived control on the considered dependent variables are lower in magnitude. Background characteristics like formal educational attainments explain some of the differences in perceived control, but most covariates are not statistically significant. Holding some school degree increases perceived control by 4.6 percent of a standard deviation. Moreover, a relatively small combined age effect seem to prevail. Apart from that, the pattern is similar to previous findings from the literature on trait determinants, where indicative individual characteristics usually have low explanatory power.

Turning to the estimates that are most important in light of the hypothesized feedback effect, past poverty experiences apparently exert some lagged influence on perceived control. In case of the binary headcount, the feedback effect amounts to 5.2 percent of a standard deviation. Though this is not a major feedback effect, it is still remarkable given the often alleged stability of personality traits in adulthood. One should bear in mind that the binary poverty status may mix the effects of those being “slightly” poor with those who have available an even lower amount of equivalent income. Thus, a more nuanced view may be obtained when additionally considering the extent of poverty by means of the deficit or Watts measure. When poverty is linearly scaled as equivalent income, the corresponding negative effect on perceived control is close to zero. When a higher emphasis is put on those individuals who suffer from a higher degree of poverty, a one percent increase in the ratio of the poverty line and an individual’s equivalence income significantly decreases perceived control by almost 13 percent of a standard deviation. This

is a fairly large effect that particularly seems to be driven by those individuals who are exerted to a comparably high degree of poverty. To give this effect size a more intuitive meaning, consider the case of a poverty line being located at 700 equivalent euros and an individual who has 350 equivalent euros at her/his disposal. For such an admittedly high degree of poverty, a one percent increase would be equivalent to a 2.40 euro decrease in disposable equivalent income. This finding indicates that the negative perception of small income decreases for poor individuals seems to be sizeable.

< Include Table 6 about here >

Though the overall picture does not substantially differ from what has been estimated for the female sample, some slight variations are apparent for male sample members, though. The results for the set of strictly exogenous covariates again is little surprising. However, human capital achievements and labor market assets, like job experience, imply some differing partial effects. Considering the compound contemporaneous and time-averaged variables, holding some secondary school degree results in almost the same average effect as for the male sample, whereas the poverty reduction due to a higher secondary degree is weaker. The latter amounts to just -1.8 and -2.1 percentage points, as opposed to -3.3 and -5.7 for females. Similarly, the average poverty reduction induced by having some vocational degree is only half that size. On the other hand, degrees from university and technical college greatly reduce the probability of living in poverty, but the latter effect is somewhat weaker than for females. Only the indicator for residence in the eastern parts of Germany increases the likelihood of living in poverty. Compared to the female respondents, a similar pattern holds for the various impacts of the contemporaneous and lagged predetermined variables. Having a full-time employment reduces the poverty risk by almost 7 percentage points and thus exhibits almost the same magnitude. The contemporaneous effect of living with a partner is somewhat weaker, whereas having at least one child increases the probability of poverty by 4.4 percentage points. This impact was not at hand in case of females in working-age. As expected, having been poor in the previous observational period

increases the current risk of poverty by 15.3 percentage points. As such, this state dependence is somewhat weaker than in case of the female sample. The estimated coefficients on the remaining lagged dependencies do not differ substantially and are relatively weak in magnitude.

< Include Tables 7 and 8 about here >

Columns II to IV of Table 6 present the estimates for the predetermined variables. As with the above discussion for the female sample, all the regressors apart from poverty status produce almost the same estimates for the three considered poverty measures (see Tables 6 to 8). Hence, it is again sufficient to discuss only the estimates presented in Table 6. Once more, a statistically significant effect of past employment on future employment is found. It seems that the relative magnitudes of the strictly exogenous variables are similar for the gender specific employment models. The compound contemporaneous and time-averaged effects for the secondary schooling degrees are slightly lower in case of the male sample. The impact of holding a university degree is somewhat higher, on the contrary. Likewise, living in the eastern part of Germany seems to have a marginally higher impact on the average risk of being unemployed than in case of females. Such minimal differences in the estimated effect sizes in the exogenous variables continue to hold for the remaining predetermined variables. They do not, however, add anything substantial to the general formation patterns of the considered entities.

The contemporaneous and lagged cross-effects provide more interesting information, however. Having children does not negatively impinge on the probability of employment, but even slightly increases the chance by 3.5 percentage points. The small negative effect of living with a partner is reversed for males. A substantial difference occurs for the association of cohabitation with a partner and the probability of having at least one child. As opposed to just 8 percentage points for females, living with a partner increases the probability of the latter by 34 percentage points. As with the discussion thus far, the state dependence effects are the most substantial determinants for all the predetermined variables involved. In case of employment and child bearing, the state dependence is slightly weaker

than for the female sample. In case of the probability of living with a partner, it is 0.5 percentage points higher. For the rightmost model that addresses perceived control, the explanatory contributions of the strictly exogenous variables are comparable to the results for females. Only the indicator for holding a secondary schooling degree provides a seizable explanation for the expression of perceived control, as it increases the score by an average of 16.5 percent of a standard deviation. The effects of the remaining estimates are again quite imprecise. There is also a relatively large state dependence in the model for perceived control, along with lagged feedbacks due to living with a partner and having at least one child. As opposed to the estimates obtained for the female sample, the lagged effect of cohabitation is quite substantial and increases the current period perceived control by 13.2 percent of a standard deviation. This effect is even stronger than the negative feedback exerted from past poverty experiences, which amounts to 4.3 percent of a standard deviation. Also in contrast to females, for whom the magnitude was comparable, this effect is statistically significant. Its interpretation in terms of monetary changes is similar to the example given above.

5.2 Structural Approach

The results for the structural model that explicitly accounts for potential feedback in the predetermined variables are shown in Table 9 for females and Table 10 for males. Regarding the possible orders of predeterminedness for the dependent variables in columns II to IV, the Lagrange multiplier test of Anderson and Kunitomo (1992) fails to reject the null hypothesis of the predeterminedness order as given by Tables 9 and 10 against overidentified (unrestricted) alternatives in case of both subsamples. This result shows that the structural model as suggested is in line with data.

< Include Table 9 about here >

Moreover, recall from Section 4 that the structural approach considers complete time paths over the whole observational timespan, which leads to a substantially lower number of observations compared to the models discussed in the

previous section. For the female sample, some of the revealed effects are different, but not in a way that is inconsistent with the previous findings. Looking at the impact of the strictly exogenous variables, apart from the general secondary schooling degree, education significantly decreases the risk of living in poverty. The indicators for holding a higher secondary schooling degree, a university degree, or a technical college degree are the only education variables the time-means of which have substantial poverty reducing effects, probably due to their role as projections of unobserved abilities. As for the partial likelihood estimates, living in the eastern part of Germany can be associated with an increase in poverty risk. The poverty reducing effect of full time job experience is slightly higher than suggested by the previous models. The results for the impact of holding a German citizenship are comparable under both models considered thus far. All further exogenous partial effects are rather negligible within the poverty equation. Regarding the contemporaneous cross-effects, the mediating role of employment seems to be more of a factor within the structural setup. As such, it may also be possible that some of the exogenous covariates additionally operate on poverty status via the employment equation. When jointly considering the contemporaneous exogenous variables and the correlated part of the employment equation, the compound effects have not changed much. The overall impacts of the human capital related characteristics have a positive influence on employment probability in large parts, whereas the impacts of graduation from university and technical college have changed their signs. By and large, there also are no dramatic changes in the exogenous and correlated model parts for the equations representing partnership and having children, though some effects are even reversed. For instance, the impact of a technical college degree on the probability of having at least one child changes from 3 percentage point to -7 percentage points, an effect that is relatively weak in magnitude though. Other effects remain almost unchanged as is exemplified by the average probability change exerted from holding a university degree to living in a partnership. Likewise maintained is the almost 30 percentage point reduction in the probability of living in a partnership for those individuals

who live in eastern Germany. Returning the attention to the contemporaneous interrelations between the model equations, the effects of having children, living with a partner, and of the individual degree of perceived control are still comparable to those found within the partial likelihood approach. The impacts of partnership and childbearing are still negative with regard to employment but have decreased, whereas the magnitude of perceived control in the partnership equation has increased to almost 7 percentage points. Furthermore, the average effect from cohabitation to having a child is up by some 3.5 percentage points compared to the previous framework.

With regard to lagged effects of the predetermined variables and poverty, the strong state dependence within the respective model parts remains for the structural model as well. In case of poverty and employment, it has decreased, whereas for the children and cohabitation sub-models, there is a slight increase in the state dependencies. The previously positive lagged effects from poverty and partnership on the probability of having a child, reverse into quite small negative effects. The explanatory associations for the model part on perceived control is still rather diffuse and at best allows to infer some significant relations with regard to age. What turns out to be the most important finding of the structural estimates, however, is that the sizeable feedback effect from previous poverty experiences to control expression seem to be confirmed. As opposed to the partial likelihood model, it has even increased to -7.6 percent of a standard deviation.

< Include Table 10 about here >

Table 10 displays the results for the structural model given the male sample. Concerning the coefficients for the poverty equation, the strictly exogenous and time-invariant effects are again comparable in magnitude. Most of the human capital related predictors lower the probability of living in poverty, as does having the German citizenship. Living in eastern Germany, on the other hand, is again negatively associated with poverty reduction. As with the female sample, the contemporaneous impact of employment is slightly higher in case of the structural model, which again may be an argument in favor of contemplating employment

as a mediator of poverty. The other effects are remarkably similar, though the direct impact of perceived control on poverty is substantially lower. The effects of the variables on employment are also in line with prior expectations. Higher educational qualifications are generally associated with higher employment probabilities. In the structural model, the effect of age on the employment probability has inverted, but still is relatively weak in magnitude. The role of the exogenous variables in the remaining model parts also follow the previous discussions in large parts. What should be noted, however, is that the structural model again suggests sizable and significant state dependence effects across the entities involved in the five equations. Those for poverty, living with a partner, and perceived control are even stronger than in the previous models, whereas the state dependence for employment and having a child have decreased. Likewise, the effects for the other lagged predetermined variables show very similar patterns to those in the previous models. The results for the employment equation are rather weak. The fact that lagged poverty slightly increases the employment probability of a given period is somewhat at odds with what one could have expected. All other lagged cross-effects are rather low with regard to their magnitudes and their interpretations. Quite remarkable though is the fact that the effect of primary interest has increased (in absolute terms) by more than 3 percentage points compared to the partial likelihood model. This finding provides further evidence that, on average, past poverty experiences seem to negatively impinge on control perception.

A final note on the estimates for the variance component σ_a , which can be directly quantified in the structural model, may be in order. As opposed to the true state dependence, the unobserved component is relatively low. This may be owed to the fact, that controlling for perceived control is expected to significantly reduce the unobserved heterogeneity that usually prevails in the compound error.

5.3 Robustness Checks

As noted above, several caveats should be considered given the estimates discussed thus far. The first issue that has been addressed was that the comparably

long observational interval is likely to affect the dynamic cross-effects and state-dependencies in the setting at hand. The other feature of the data that should be taken into account is that the time-averaged model parts fluctuate somewhat more than the other derived effects. This may be an artifact of the relatively low number of waves that can be used for the considered models. This section addresses both points in successive order.

< Include Tables 11 and 12 about here >

Regarding the potential problems arising from the quite distant observational points, Tables 11 and 12 provide partial likelihood estimates for the annually available waves from 1994 to 1996, where the 1994 wave acts as initial period. Due to the nature of the partial likelihood approach, only the time-averaged effects are likely to suffer from this even shorter timespan. All other effects may provide a viable comparison to the main results. As such, we will not focus on the differences in the strictly exogenous variables and the correlated model parts here. Regarding the cross-dependencies for the female sample, one finds that relatively minor changes occur for the contemporaneous impacts of the predetermined variables. The risk-reducing effect of employment on poverty has decreased by an absolute margin of 2 percentage points, whereas the other contemporaneous effects in the poverty model remain remarkably stable compared to the original model. For the employment and childbearing model, only the effects of having a child and living with a partner have become weaker. Substantial changes occur for the state dependencies in all five models. This result is likely to arise due to the large discrepancy between the unit period and the observational interval in the original models. Apart from few exceptions, the lagged cross effects seem to be less influenced by this issue. In particular, the feedback effect from poverty to perceived control is only 0.7 percentage points weaker in absolute terms. This indicates a sufficient degree of robustness for the main results on this association.

For the male sample, the picture is quite similar. Some of the contemporaneous cross-effects have changed, but not in a substantial way. Again, the employment effect on poverty is somewhat weaker. Moreover, the effect of living with a part-

ner has reversed, but only amounts to a 4 percentage point change regarding the probability of living in poverty. For the remaining cross-effects, the changes are negligible. As opposed to the female sample, the vast increases in the state dependencies pertain to that of having at least one child, only. All other path dependencies are only moderately increased. Many of the lagged cross-dependencies that are subject to some changes are below 5 percentage points in both models and thus only provide quite unsystematic findings. The negative feedback from past poverty to perceived control has more than doubled, on the other hand. If one sees the former results on this effect as a lower bound estimate, this result at least does not jeopardize the hypothesis of a non-zero feedback on perceived control.

< Include Tables 13 and 14 about here >

Looking at the results in Tables 13 and 14, some robustness checks on the potential effect of the rather low number of waves in the models thus far are provided. The employed waves comprise those from 1994 to 2010, with the 1994 wave again representing the initial period. As perceived control is not available on an annual basis for this timespan, the checks only comprise the first four equations of the original model. Regarding the female respondents, the differences in the correlated model parts are only minor when comparing the long panel with original model. The low number of within-individual observations seems to be more problematic for the contemporaneous exogenous variables, as the differences are most substantial in these model parts. The contemporaneous cross-effects for the predetermined variables are roughly in line with the original ones and those for the first robustness check. Fortunately, the same seems to hold for the lagged model parts. As with the model for the waves from 1994 to 1996, the state dependencies have substantially increased as a result of the annual observational interval. However, the lagged cross-effects seem to remain comparable in large parts. It may be cautiously concluded that this would also translate to the lagged feedback on perceived control, given it would have been available. For the male sample the same pattern of changes seems to apply. Again, there are quite substantial

differences within the exogenous and correlated model parts, as well as increases in the magnitudes of the state dependencies within the four equations. On the other hand, the contemporaneous and lagged crosseffects are rather in line with the previously examined model specifications.

The strong changes in the state dependencies provide evidence in favor of the points previously made on the mixing of different spell types. Especially for poverty and employment in the 5-year observational interval, those individuals for whom the two points in time are part of a continuing spell and those individuals who have one or more interrupted spell(s) obviously mix up in the original estimate.

6 Conclusion

Summarizing the previous findings, some rather robust results that establish the existence of a lagged feedback from poverty experiences to perceived control are provided by the presented dynamic panel estimates. Compared to early interventional studies (see, e.g., Almlund et al., 2011), most of the retrieved impacts on perceived control are comparably low. They are, however, sufficiently large in order to claim that the assumption of complete invariance of control-perception in adulthood is inappropriate in some cases. Accordingly, the negative effects of past poverty experiences range from 4 to 10 percent of a standard deviation of perceived control. Referring to the trait formation literature, the results thus far are largely in line with previous findings that advocate small impacts of trigger events in adulthood on the stability of personality traits like perceived control (see, e.g., Cobb-Clark and Schurer, 2013). As opposed to the setting at hand, these are one time occurrences, however. As such, this pattern seems to be altered slightly when the persistence of the event is considered. These findings support the hypothesis that there can be sizeable changes in attitudes when an individual experiences certain long lasting environmental changes.

Whether this result also shows that personality traits in general are susceptible to similar changes is not resolved by the empirical findings presented here.

Considering the inverse causal association, control attitudes do not provide any substantial information on the probability of slipping into poverty in the first place. The poverty status seems to primarily result from large state dependencies in poverty and employment, as well as from the corresponding cross-effects. Taken together, the results suggest that, apart from the alleged channel via perceived control, poverty experiences are further associated with processes of depreciation of human capital, demoralization, and incentive reductions. These mediators seem to jointly increase the probability that individuals who become poor will remain so for extended periods.

With respect to the other entities that are incorporated into the model frameworks, some additional interesting insights can be obtained. From a methodological perspective, the existence of feedback effects across equations on the future values of the predetermined variables makes the use of traditional random or fixed effects models, which are based on the strict exogeneity assumption, questionable with respect to the current and related settings. Based on the framework of Wooldridge (2000), the empirical analysis at hand draws on dynamic models that explicitly allows for such feedbacks. Previous estimation results suggest that feedback effects indeed prevail in lots of panel data settings, for instance in case of low (wage-)incomes and employment (see Stewart, 2007). Given the estimates provided here for German data, however, low equivalent incomes in the previous period do not impinge on the employment probability of the next period. There are indications for feedbacks on other entities, though. For instance, there is evidence that poverty affects household constitution, though in different directions for males and females. However, household and family formation are quite complex decision problems and the results presented here are indicative at best. The further reflected state dependence effects have been found in previous studies as well. For instance in case of (un)employment dynamics Arulampalam et al. (2000) also show large degrees of state dependence. As for the present results, it has been additionally shown that the extent of state dependence is subject to the chosen observational interval in large parts.

References

- AASSVE, A., S. BURGESS, A. CHESHER, AND C. PROPPER (2002): “Transitions From Home to Marriage of Young Americans,” *Journal of Applied Econometrics*, 17(1), 1–23.
- AASSVE, A., S. BURGESS, M. DICKSON, AND C. PROPPER (2006a): “Modelling Poverty by not Modelling Poverty: An Application of a Simultaneous Hazard Approach to the UK,” CASEpaper 106, Centre for Analysis of Social Exclusion - London School of Economics, London.
- AASSVE, A., S. BURGESS, C. PROPPER, AND M. DICKSON (2006b): “Employment, Family Union and Childbearing Decisions in Great Britain,” *Journal of the Royal Statistical Society Series A*, 169(4), 781–804.
- AINSLIE, G. (1975): “Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control,” *Psychological Bulletin*, 82(4), 463–496.
- (1991): “Derivation of “Rational” Economic Behavior from Hyperbolic Discount Curves,” *American Economic Review*, 81(2), 334–340.
- ALMLUND, M., A. L. DUCKWORTH, J. HECKMAN, AND T. KAUTZ (2011): “Personality Psychology and Economics,” in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. J. Machin, and L. Woessmann, vol. 4, chap. 1, pp. 1–158. Elsevier B.V., Amsterdam: North-Holland.
- ANDERSON, T. W., AND N. KUNITOMO (1992): “Tests of Overidentification and Predeterminedness in Simultaneous Equation Models,” *Journal of Econometrics*, 54, 49–78.
- ANDERSON, T. W., AND H. RUBIN (1956): “Statistical Inference in Factor Analysis,” in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman, pp. 111–150. Berkeley - University of California Press, Berkeley, California.

- APPADURAI, A. (2004): “The Capacity to Aspire,” in *Culture and Public Action*, ed. by V. Rao, and M. Walton, pp. 59–84. The International Bank for Reconstruction and Development, The World Bank, Washington, DC.
- ARELLANO, M., AND B. HONORÉ (2001): “Panel Data Models: Some Recent Developments,” in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. E. Leamer, vol. 5, chap. 76, pp. 3229–3296. Elsevier B.V., Amsterdam, North Holland.
- ARULAMPALAM, W., A. L. BOOTH, AND M. P. TAYLOR (2000): “Unemployment Persistence,” *Oxford Economic Papers*, 52(1), 24–50.
- ATKINSON, A. B. (1970): “On the Measurement of Inequality,” *Journal of Economic Theory*, 2(3), 244–263.
- ATKINSON, A. B., L. RAINWATER, AND T. M. SMEEDING (1995): *Income Distribution in OECD Countries*. OECD, Paris.
- BALTAGI, B., AND S. SONG (2006): “Unbalanced Panel Data: A Survey,” *Statistical Papers*, 47(4), 493–523.
- BARRON, A. R., AND C.-H. SHEU (1991): “Approximation of Density Functions by Sequences of Exponential Families,” *The Annals of Statistics*, 19(3), 1347–1369.
- BECKER, G. S. (1993): *A Treatise on the Family*. Harvard University Press.
- BERNHEIM, B. D., D. RAY, AND S. YELTIKIN (2013): “Poverty and Self-Control,” NBER Working Papers 18742, National Bureau of Economic Research (NBER), Cambridge, MA.
- BIEWEN, M. (2004): “Measuring State Dependence in the Individual Poverty Status: Are There Feedback Effects to Employment Decisions and Household Composition?,” IZA Discussion Paper Series 1138, Institute for the Study of Labor (IZA), Bonn.

- BROWNING, M., P.-A. CHIAPPORI, AND Y. WEISS (2011): *Economics of the Family*. Cambridge University Press, Cambridge.
- BURDETT, K., AND M. G. COLES (1999): “Long-Term Partnership Formation: Marriage and Employment,” *Economic Journal*, 109(456), F307–34.
- BUTLER, J. S., AND R. MOFFITT (1982): “A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model,” *Econometrica*, 50(3), 761–764.
- CHAMBERLAIN, G. (1982): “Multivariate Regression Models for Panel Data,” *Journal of Econometrics*, 18(1), 5–46.
- CHEN, X. (2007): “Large Sample Sieve Estimation of Semi-Nonparametric Models,” in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. E. Leamer, vol. 6B, chap. 76, pp. 5549–5632. Elsevier B.V., Amsterdam, North Holland.
- COBB-CLARK, D. A., AND S. SCHURER (2013): “Two Economists’ Mursings on the Stability of Locus of Control,” *Economic Journal*, 123(8), 358–400.
- COSTELLO, A. B., AND J. W. OSBORNE (2005): “Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis,” *Practical Assessment, Research & Evaluation*, 10(7).
- COWELL, F. A. (2011): *Measuring Inequality*. Oxford University Press, Oxford, NY.
- DAGSVIK, J. K., S. STROM, AND Z. JIA (2006): “Utility of income as a random function: Behavioral characterization and empirical evidence,” *Mathematical Social Sciences*, 51(1), 23–57.
- DALTON, P. S., S. GHOSAL, AND A. MANI (2013): “Poverty and Aspirations Failure,” Discussion paper, Tilburg University and University of Warwick, mimeo.
- DAVIDSON, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*. Oxford University Press.

- ENGLE, R. F., D. F. HENDRY, AND J.-F. RICHARD (1983): “Exogeneity,” *Econometrica*, 51(2), 277–304.
- FOSTER, J., AND A. SEN (1997): *On Economic Inequality*. Oxford University Press, New York.
- FOSTER, J., S. SETH, M. LOKSHIN, AND Z. SAJAIA (2013): *A Unified Approach to Measuring Poverty and Inequality - Theory and Practice*. The World Bank, Washington, DC.
- FULEKY, P. (2012): “On the Choice of the Unit Period in Time Series Models,” *Applied Economics Letters*, 19(12), 1179–1182.
- GENICOT, G., AND D. RAY (2013): “Aspirations and Inequality,” Discussion paper, New York University, mimeo.
- HAUSHOFER, J., AND E. FEHR (2014): “On the Psychology of Poverty,” *Science*, 344(6186), 862–867.
- HECKMAN, J., L. LOCHNER, AND P. TODD (2006): “Fifty Years of Mincer Earinings Regression,” in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–451. Elsevier B.V., Amsterdam: North-Holland.
- HECKMAN, J., J. STIXRUD, AND S. URZUA (2006): “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,” *Journal of Labor Economics*, 24(3), 411–482.
- HEINECK, G., AND S. ANGER (2010): “The Returns to Cognitive Abilities and Personality Traits in Germany,” *Labour Economics*, 17(3), 535–546.
- HOJMAN, D., AND F. KAST (2009): “On the Measurement of Poverty Dynamics,” Working paper series, Harvard University, John F. Kennedy School of Government.
- HÄRDLE, W. (1995): *Applied Nonparametric Regression*. Cambridge University Press, New York, NY.

- HSIAO, C. (2003): *Analysis of Panel Data*. Cambridge University Press., New York, NY.
- JUDD, K. (1998): *Numerical Methods in Economics*. MIT Press, Cambridge, Massachusetts.
- KILIAN, L., AND C. VEGA (2011): “Do Energy Prices Respond to U.S. Macroeconomic News? A Test of the Hypothesis of Predetermined Energy Prices,” *The Review of Economics and Statistics*, 93(2), 660–671.
- LAIBSON, D. I. (1997): “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 112, 443–477.
- LAMBERT, P. J. (2001): *The Distribution and Redistribution of Income*. Manchester University Press., New York, NY.
- MADDALLA, G. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press., New York, NY.
- MCCULLAGH, P., AND J. A. NELDER (1989): *Generalized Linear Models*. Chapman and Hall, London, United Kingdom.
- MILLIMET, D. L., AND I. K. MCDONOUGH (2013): “Dynamic Panel Data Models with Irregular Spacing: With Applications to Early Childhood Development,” IZA Discussion Papers, Institute for the Study of Labor (IZA).
- MUELLER, G., AND E. PLUG (2006): “Estimating the Effect of Personality on Male and Female Earnings,” *Industrial and Labor Relations Review*, 60(1), 3–22.
- MULAİK, S. A. (2009): *Linear Causal Modelling with Structural Equations*. Chapman & Hall/CRC Press, Boca Raton, FL.
- MUNDLAK, Y. (1978): “On the Pooling of Time Series and Cross Section Data,” *Econometrica*, 46(1), 69–85.

- MURAVEN, M., AND R. F. BAUMEISTER (2000): “Self-Regulation and Depletion of Limited Resources: Does Self-Control resemble a muscle,” *Psychological Bulletin*, 126(2), 247–257.
- NEYMAN, J., AND E. L. SCOTT (1948): “Consistent Estimates Based on Partially Consistent Observations,” *Econometrica*, 16(1), 1–32.
- RENCHEK, A. (2004): *Methods of Multivariate Analysis*. Wiley Interscience.
- ROBERTS, B. W., W. VIECHTBAUER, AND K. E. WALTON (2006): “Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies,” *Psychological Bulletin*, 132(1), 1–25.
- ROTTER, J. (1966): *Generalized Expectancies for Internal versus External Control of Reinforcement*. American Psychological Association, Washington, DC.
- SEN, A. (1976): “Poverty: An Ordinal Approach to Measurement,” *Econometrica*, 44(2), 219–231.
- (1982): “Equality of What?,” in *Choice, Welfare and Measurement*, ed. by A. Sen, chap. 16, pp. 353–369. Blackwell, Oxford.
- SHER, J. (1977): *Education in Rural America*. Westview: Boulder.
- SKRONDAL, A., AND S. RABE-HESKETH (2004): *Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC, Boca Raton, FA.
- SPADY, R. H. (2006): “Identification and Estimation of Latent Attitudes and their Behavioral Implications,” CEMMAP Working Paper CWP12/06, Centre for Micro Data Methods and Practice (CEMMAP), London.
- (2007): “Semiparametric Methods for the Measurement of Latent Attitudes and the Estimation of their Behavioral Consequences,” CEMMAP Working Paper CWP26/07, Centre for Micro Data Methods and Practice (CEMMAP), London.

- STEVENS, A. H. (1999): “Climbing out of Poverty, Falling Back in: Measuring the Persistence of Poverty Over Multiple Spells,” *Journal of Human Resources*, 34(3), 557–588.
- STEWART, M. B. (2007): “The Interrelated Dynamics of Unemployment and Low-Wage Employment,” *Journal of Applied Econometrics*, 22(3), 511–531.
- TANGNEY, J., R. BAUMEISTER, AND M. BOONE (2004): “High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success,” *Journal of Personality*, 72(2), 271–322.
- THALER, R. H., AND S. BENARTZI (2004): “Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving,” *Journal of Political Economy*, 112(S1), 164–S187.
- THEIL, H. (1967): *Economics and Information Theory*. Elsevier, Amsterdam, North Holland.
- THIEL, H., AND S. THOMSEN (2013): “Noncognitive Skills in Economics: Models, Measurement, and Empirical Evidence,” *Research in Economics*, 67(2), 189–214.
- THOMPSON, B. (2004): *Exploratory and Confirmatory Factor Analysis - Understanding Concepts and Applications*. American Psychological Association, Washington D.C.
- WAGNER, G. G., J. R. FRICK, AND J. SCHUPP (2007): “The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements,” *Schmollers Jahrbuch*, 127, 139–169.
- WATTS, H. (1968): “An Economic Definition of Poverty,” in *On Understanding Poverty: Perspectives from the Social Sciences*, ed. by D. P. Moynihan. Basic Books, New York, NY.
- WHITE, H. (1982): “Maximum Likelihood Estimation of Misspecified Models,” *Econometrica*, 50(1), 1–25.

- WOOLDRIDGE, J. M. (2000): "A Framework for Estimating Dynamic, Unobserved Effects Panel Data Models with Possible Feedback to Future Explanatory Variables," *Economic Letters*, 68, 245–250.
- (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Hong Kong.
- (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20(1), 39–54.
- ZHENG, B. (1992): "An Axiomatic Characterization of the Watts Poverty Index," *Economic Letters*, 42, 81–86.
- (1994): "Can a Poverty Index be Both Relative and Absolute," *Econometrica*, 62(6), 1453–1458.
- (1997): "Aggregate Poverty Measures," *Journal of Economic Surveys*, 11(2), 123–162.

Tables and Figures

Table 1: Sample Descriptives (Wave 2005)

	Non-Poverty		Poverty		Mean-Diff.
	Mean	Stand. Dev.	Mean	Stand. Dev.	<i>p</i> -value ^b
<i>Female</i>					
Employed (D)	0.330	0.470	0.099	0.299	0.000
Perceived Control ^a	0.050	0.984	-0.366	1.054	0.000
Some School (D)	0.667	0.471	0.671	0.470	0.779
Higher Secondary (D)	0.297	0.457	0.235	0.424	0.000
University (D)	0.125	0.331	0.037	0.190	0.000
Job Experience (Full Time)	11.373	10.760	7.731	9.370	0.000
Age	42.137	13.170	38.319	13.658	0.000
East German (D)	0.205	0.403	0.319	0.466	0.000
Child(ren) in HH (D)	0.411	0.492	0.524	0.500	0.000
Living with Partner (D)	0.635	0.481	0.378	0.485	0.000
Sample Size (N)	7,692		1,120		
<i>Male</i>					
Employed (D)	0.716	0.451	0.222	0.416	0.000
Perceived Control ^a	0.048	0.970	-0.481	1.106	0.000
Some School (D)	0.659	0.474	0.641	0.480	0.303
Higher Secondary (D)	0.301	0.459	0.258	0.438	0.012
University (D)	0.152	0.359	0.035	0.185	0.000
Job Experience (Full Time)	19.091	12.768	13.918	12.416	0.000
Age	42.256	13.335	38.647	13.865	0.000
East German (D)	0.206	0.404	0.368	0.483	0.000
Child(ren) in HH (D)	0.395	0.489	0.465	0.499	0.000
Living with Partner (D)	0.608	0.488	0.414	0.493	0.000
Sample Size (N)	7,545		939		

Dummy variables are indicated by (D).

^a Standardized raw scores.

^b Two-sample equality of mean t-test.

Table 2: Variable Types and Corresponding Link Functions $g(v)$

Variables	Link Type	Range of $g(v)$	$g(v)$
Poverty Metrics			
Headcount	binary	$\{0, 1\}$	$\Phi(v)$
Poverty Deficit	corner solution	$(0, +\infty)$	$\Phi(v/\sigma)(v + \sigma \left[\frac{\phi(v/\sigma)}{\Phi(v/\sigma)} \right])$
Watts	corner solution	$(0, +\infty)$	$\Phi(v/\sigma)(v + \sigma \left[\frac{\phi(v/\sigma)}{\Phi(v/\sigma)} \right])$
(Potentially) Predetermined Variables			
≥ 1 child in HH	binary	$\{0, 1\}$	$\Phi(v)$
Living with partner	binary	$\{0, 1\}$	$\Phi(v)$
Employment (full time)	binary	$\{0, 1\}$	$\Phi(v)$
Perceived Control	identity	$(-\infty, +\infty)$	v

Table 3: Average Partial Effects for Headcount (Female) – Partial Likelihood Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.069***	–	–	–	–
Children	0.003**	-0.220**	–	–	–
Partner in HH	-0.065***	-0.029***	0.081***	–	–
Perceived Control (PC)	-0.012***	0.011***	-0.001	0.010***	–
Lagged Variables					
Poverty St.	0.189***	0.009**	0.059***	-0.042***	-0.052
Employment St.	0.017***	0.309***	0.024**	0.001	-0.054**
Children	0.016***	0.127***	0.502***	-0.011***	-0.024
Partner in HH	0.011***	-0.015***	0.176***	0.207***	-0.013
Perceived Control (PC)	-0.007**	0.003**	0.007**	0.006**	0.369***
Strictly Exogenous Variables					
Some School (D)	-0.020***	0.028**	0.113***	-0.056***	0.046*
Higher Secondary (D)	-0.033***	-0.062***	0.082***	-0.080***	-0.130
Some Voc. Train. (D)	-0.038**	0.070**	0.019**	0.008**	-0.040*
University (D)	0.002	0.333***	0.156***	0.084***	0.068*
Tech-Coll. (D)	-0.010	0.264***	0.080***	0.022**	0.093
Job Exp. (Full T.)	0.001*	-0.006***	-0.001**	0.004***	0.002
Age	-0.004***	-0.008**	-0.011***	-0.004***	-0.026***
East German (D)	0.067***	0.050**	-0.041***	-0.138***	-0.050
German (D)	-0.035**	-0.024*	0.194***	0.071***	-0.076
Time Averages					
Some School (D)	-0.055***	0.021**	-0.105**	0.064***	0.087*
Higher Secondary (D)	-0.057***	0.012***	-0.084***	0.077***	0.190
Some Voc. Train. (D)	0.001	0.004**	0.013**	0.014**	0.047
University (D)	-0.060***	-0.158***	-0.114***	-0.076***	0.156*
Tech-Coll. (D)	-0.041***	-0.155***	-0.053***	-0.003	-0.025
Job Exp. (Full T.)	-0.002***	0.019***	-0.003***	-0.005***	0.006*
Age	-0.005***	-0.007**	-0.002***	0.003***	0.018***
East German (D)	0.001	-0.073***	0.004	-0.135***	-0.064
German (D)	-0.001	0.022	-0.245**	-0.121***	-0.025*
<i>N</i>	8,954	9,079	12,069	9,067	8,489

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Average Partial Effects for Poverty Deficit (Female) – Partial Likelihood Approach (1995 – 2010, Sampling Period = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-199.414***	–	–	–	–
Children	-2.240	-0.219***	–	–	–
Partner in HH	-160.355***	-0.028***	0.080***	–	–
Perceived Control (PC)	-18.239***	0.011***	-0.002*	0.010***	–
Lagged Variables					
Poverty St.	0.979***	0.000***	0.000***	-0.000***	-0.000
Employment St.	20.015	0.309***	0.022***	0.002	-0.052**
Children	34.303**	0.127***	0.502***	-0.010***	-0.024
Partner in HH	3.953	-0.015***	0.174***	0.207***	-0.011
Perceived Control (PC)	-17.853***	0.003***	0.006***	0.006***	0.369***
Strictly Exogenous Variables					
Some School (D)	-39.425	0.028**	0.115***	-0.059***	0.041
Higher Secondary (D)	-76.730	-0.061***	0.083***	-0.083***	-0.019
Some Voc. Train. (D)	-59.488**	0.070**	0.021***	0.008**	0.039*
University (D)	14.693	0.333***	0.155***	0.084***	-0.067
Tech-Coll. (D)	18.197	0.264***	0.081***	0.021**	0.092
Job Exp. (Full T.)	2.855	-0.005***	-0.001**	0.004***	0.002
Age	6.456***	-0.008**	-0.010***	-0.004***	-0.026***
East German (D)	86.869**	0.050**	-0.039***	-0.139***	-0.052
German (D)	-45.067	-0.025*	0.197***	0.067***	-0.077
Time Averages					
Some School (D)	-73.328	0.021*	-0.108***	0.067***	0.093
Higher Secondary (D)	-62.061	0.013***	-0.086***	0.079***	0.196*
Some Voc. Train. (D)	-21.336	0.004**	0.011**	0.015**	0.048*
University (D)	-126.505*	-0.158***	-0.115***	-0.076***	0.156*
Tech-Coll. (D)	-124.519*	-0.155***	-0.056***	-0.001	-0.023
Job Exp. (Full T.)	-4.704	0.019***	-0.003***	-0.005***	0.005*
Age	-8.786***	-0.007**	-0.002***	0.003***	0.017***
East German (D)	18.799	-0.073***	0.004	0.135***	-0.064
German (D)	-16.633	0.022	-0.249***	-0.116***	0.251*
<i>N</i>	8,954	9,079	12,069	9,067	8,489

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Average Partial Effects for Watts Measure (Female) – Partial Likelihood Approach (1995 – 2010, Sampling Period = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.408***	–	–	–	–
Children	0.018	-.219***	–	–	–
Partner in HH	-0.327***	-.029***	0.080***	–	–
Perceived Control (PC)	-0.034***	.011***	-0.002*	0.010***	–
Lagged Variables					
Poverty St.	0.718***	0.036***	0.109***	-0.059***	-0.129*
Employment St.	0.025	0.309***	0.021***	0.004**	-0.053**
Children	0.055**	0.127***	0.503***	-0.011***	-0.024
Partner in HH	-0.007	-0.015***	0.173***	0.208***	-0.011
Perceived Control (PC)	-0.042***	0.003***	0.006***	0.007***	0.370***
Strictly Exogenous Variables					
Some School (D)	-0.095	0.028**	0.115***	-0.058***	0.039
Higher Secondary (D)	-0.172*	-0.061***	0.083***	-0.082***	-0.020
Some Voc. Train. (D)	-0.125**	0.070***	0.021***	0.008**	0.039*
University (D)	0.023	0.333***	0.156***	0.084***	-0.067
Tech-Coll. (D)	0.020	0.264***	0.081***	0.021**	0.092
Job Exp. (Full T.)	0.008	-0.006***	-0.001**	0.003***	0.002
Age	0.013***	-0.008***	-0.010***	-0.004***	-0.026***
East German (D)	0.168*	0.049***	-0.039***	-0.139***	-0.051
German (D)	-0.118	-0.025*	0.197***	0.066***	-0.077
Time Averages					
Some School (D)	-0.135	0.021*	-0.109***	0.068***	0.093
Higher Secondary (D)	-0.121	0.128***	-0.087***	0.080***	0.196*
Some Voc. Train. (D)	-0.049	0.009**	0.010**	0.015**	0.048*
University (D)	-0.247*	-0.158***	-0.115***	-0.075***	0.156*
Tech-Coll. (D)	-0.247*	-0.156***	-0.056***	-0.001	-0.023
Job Exp. (Full T.)	-0.011	0.019***	-0.003***	-0.005***	0.005*
Age	-0.018***	-0.007***	-0.002***	0.003***	0.017***
East German (D)	0.059	-0.073***	0.005	0.134***	-0.064
German (D)	-0.015	0.023*	-0.249***	-0.115***	0.251*
<i>N</i>	8,954	9,079	12,069	9,067	8,489

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Average Partial Effects for Headcount (Male) – Partial Likelihood Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.068***	–	–	–	–
Children	0.044***	0.035***	–	–	–
Partner in HH	-0.032***	0.045***	0.340***	–	–
Perceived Control (PC)	-0.007***	0.022***	-0.002***	0.008***	–
Lagged Variables					
Poverty St.	0.153***	0.001	0.041***	0.002	-0.043*
Employment St.	0.005***	0.170***	0.111***	0.026***	0.017
Children	0.014***	0.036***	0.426***	-0.014***	-0.079***
Partner in HH	0.003	0.011***	0.088***	0.212***	0.132***
Perceived Control (PC)	-0.002***	0.004***	0.004***	0.009***	0.367***
Strictly Exogenous Variables					
Some School (D)	0.020***	0.069***	0.004	0.018**	0.165*
Higher Secondary (D)	-0.018***	0.120***	0.069***	0.056***	0.079
Some Voc. Train. (D)	-0.009***	0.053***	0.047**	0.029**	0.080*
University (D)	-0.042***	0.193***	0.058***	0.093***	0.035
Tech-Coll. (D)	-0.068***	0.186***	0.031***	0.094**	-0.007
Job Exp. (Full T.)	0.000	0.015***	-0.009**	-0.002***	0.004
Age	0.005***	-0.022***	-0.006***	0.000	-0.036***
East German (D)	0.046***	-0.059***	0.029***	-0.097***	-0.075
German (D)	-0.026**	-0.124***	0.077***	0.062***	-0.157
Time Averages					
Some School (D)	-0.057***	-0.029**	0.001	-0.021**	-0.045
Higher Secondary (D)	-0.021***	-0.108***	-0.081***	-0.046***	0.076
Some Voc. Train. (D)	-0.011***	0.040***	-0.021***	-0.014***	-0.006
University (D)	-0.038***	-0.013	-0.039***	-0.083***	0.044
Tech-Coll. (D)	0.041***	-0.131***	-0.010	-0.069***	0.030
Job Exp. (Full T.)	-0.003***	0.002***	0.002***	0.005***	0.002
Age	-0.003***	-0.005***	-0.003***	-0.003***	0.025***
East German (D)	-0.004	0.000	-0.082***	0.096***	-0.015
German (D)	0.006	0.125***	-0.136**	-0.120***	0.296
<i>N</i>	8,378	8,491	11,217	8,479	7,916

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Average Partial Effects for Poverty Deficit (Male) – Partial Likelihood Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-223.944***	–	–	–	–
Children	77.177***	0.035***	–	–	–
Partner in HH	-80.747***	0.045***	0.341***	–	–
Perceived Control (PC)	-16.007***	0.022***	-0.002***	0.008***	–
Lagged Variables					
Poverty St.	0.839***	0.000***	0.000***	0.000**	-0.000
Employment St.	1.677	0.171***	0.109***	0.026***	0.021
Children	35.359**	0.036***	0.427***	-0.014***	-0.079***
Partner in HH	-1.653	0.011***	0.089***	0.213***	0.132***
Perceived Control (PC)	-4.629	0.005***	0.004***	0.009***	0.367***
Strictly Exogenous Variables					
Some School (D)	11.1236	0.069***	0.003	0.018**	0.164*
Higher Secondary (D)	-51.2933	0.120***	0.069***	0.056***	0.077
Some Voc. Train. (D)	-11.2031	0.053***	0.047***	0.029**	0.079*
University (D)	-80.1962	0.193***	0.059***	0.093***	0.032
Tech-Coll. (D)	-205.191***	0.186***	0.033***	0.094**	-0.009
Job Exp. (Full T.)	.542645	0.015***	-0.009**	-0.002***	0.004
Age	9.30198***	-0.022***	-0.005***	0.000	-0.036***
East German (D)	41.6851	-0.059***	0.029***	-0.097***	-0.075
German (D)	-83.344	-0.119***	0.077***	0.062***	-0.157
Time Averages					
Some School (D)	-69.969*	-0.029**	0.001	-0.021**	-0.043
Higher Secondary (D)	-7.720	-0.108***	-0.083***	-0.046***	0.079
Some Voc. Train. (D)	-29.148	0.040***	-0.022***	-0.014***	-0.005
University (D)	-112.633*	-0.013	-0.041***	-0.083***	0.048
Tech-Coll. (D)	77.097	-0.130***	-0.011	-0.069***	0.033
Job Exp. (Full T.)	-5.816*	0.002***	0.002***	0.005***	0.002
Age	-5.386**	-0.005***	-0.003***	-0.003***	0.025***
East German (D)	31.783	0.000	-0.082***	0.096***	-0.015
German (D)	42.788	0.122***	-0.136***	-0.120***	0.296
<i>N</i>	8,378	8,491	11,217	8,479	7,916

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Average Partial Effects for Watts Measure (Male) – Partial Likelihood Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.418***	–	–	–	–
Children	0.156***	0.035***	–	–	–
Partner in HH	-0.157***	0.045***	0.341***	–	–
Perceived Control (PC)	-0.031***	0.022***	-0.002***	0.008***	–
Lagged Variables					
Poverty St.	0.608***	0.044***	0.080***	0.022**	-0.069
Employment St.	-0.021	0.171***	0.109***	0.026***	0.020
Children	0.062**	0.036***	0.427***	-0.014***	-0.080***
Partner in HH	-0.002	0.011***	0.089***	0.213***	0.132***
Perceived Control (PC)	-0.011	0.005***	0.004***	0.009***	0.367***
Strictly Exogenous Variables					
Some School (D)	0.002	0.069***	0.003	0.018**	0.164*
Higher Secondary (D)	-0.094	0.120***	0.069***	0.056***	0.077
Some Voc. Train. (D)	-0.023	0.053***	0.047***	0.029**	0.079*
University (D)	-0.147	0.193***	0.059***	0.093***	0.032
Tech-Coll. (D)	-0.377***	0.186***	0.033***	0.094**	-0.009
Job Exp. (Full T.)	0.001	0.015***	-0.009**	-0.002***	0.004
Age	0.017***	-0.022***	-0.005***	0.000	-0.036***
East German (D)	0.061	-0.061***	0.029***	-0.098***	-0.074
German (D)	-0.193	-0.119***	0.076***	0.063***	-0.149
Time Averages					
Some School (D)	-0.123*	-0.029**	0.001	-0.020**	-0.043
Higher Secondary (D)	-0.029	-0.108***	-0.083***	-0.045***	0.079
Some Voc. Train. (D)	-0.053	0.040***	-0.022***	-0.014***	-0.005
University (D)	-0.222*	-0.013	-0.041***	-0.083***	0.048
Tech-Coll. (D)	0.127	-0.129***	-0.011	-0.068***	0.033
Job Exp. (Full T.)	-0.012*	0.002***	0.002***	0.005***	0.002
Age	-0.009**	-0.005***	-0.003***	-0.003***	0.025***
East German (D)	0.083	0.001	-0.082***	0.096***	-0.016
German (D)	0.109	0.121***	-0.135***	-0.122***	0.289
<i>N</i>	8,378	8,491	11,217	8,479	7,916

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Average Partial Effects for Headcount (Female) – Structural Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.098***	–	–	–	–
Children	0.046**	-0.054***	–	–	–
Partner in HH	-0.011**	-0.033**	0.115***	–	–
Perceived Control (PC)	-0.079**	0.046*	0.059**	0.069***	–
Lagged Variables					
Poverty St.	0.143**	0.018**	-0.022**	0.006*	-0.076**
Employment St.	0.003**	0.101*	0.106**	0.010**	-0.015**
Children	0.014***	-0.098*	0.786**	-0.014***	-0.008*
Partner in HH	0.005*	-0.011**	-0.013*	0.212***	0.018
Perceived Control (PC)	-0.007**	0.004*	-0.059*	0.010*	0.167**
Strictly Exogenous Variables					
Some School (D)	0.020**	0.039***	0.020**	0.085	0.054*
Higher Secondary (D)	-0.021*	0.043*	0.021***	-0.058***	-0.213
Some Voc. Train. (D)	-0.009***	0.126***	0.121*	0.014*	-0.025*
University (D)	-0.052**	0.148***	0.109**	0.058***	0.071*
Tech-Coll. (D)	-0.068*	0.137*	-0.013**	0.009**	0.019
Job Exp. (Full T.)	0.005**	0.177**	0.030*	0.033**	0.002*
Age	-0.005**	-0.086***	-0.049*	0.009**	-0.021**
East German (D)	0.055***	0.059**	-0.025*	-0.171***	-0.055
German (D)	-0.026*	0.061*	0.163**	0.026*	-0.038*
Time Averages					
Some School (D)	0.013**	0.011**	-0.105**	0.016**	0.124*
Higher Secondary (D)	-0.036**	0.001*	-0.084***	0.067**	0.212
Some Voc. Train. (D)	0.007	0.013**	0.013**	0.012**	0.066
University (D)	-0.046***	-0.158***	-0.114***	-0.043***	0.135*
Tech-Coll. (D)	-0.024***	-0.155***	-0.053***	-0.007*	-0.023
Job Exp. (Full T.)	-0.012**	0.019***	-0.003***	-0.003*	-0.001*
Age	-0.011***	-0.007**	-0.002***	0.001**	0.065***
East German (D)	0.003*	-0.073***	0.004	-0.117**	-0.061
German (D)	-0.016**	0.022	-0.245**	-0.093***	-0.021*
σ_a	0.134***	0.221**	0.245**	0.121***	0.251**
N					1,489

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Average Partial Effects for Headcount (Male) – Structural Approach (1995 – 2010, Sampling Interval = 5 yrs.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.081**	–	–	–	–
Children	0.065***	0.048**	–	–	–
Partner in HH	-0.019*	0.019**	0.567**	–	–
Perceived Control (PC)	-0.093***	0.031***	0.008**	0.006***	–
Lagged Variables					
Poverty St.	0.251***	0.008**	0.037**	0.007***	-0.081***
Employment St.	0.012**	0.127***	0.109**	0.031**	0.025**
Children	0.021*	0.089*	0.368***	-0.021***	-0.001*
Partner in HH	0.009**	-0.043**	-0.001*	0.332**	0.005*
Perceived Control (PC)	-0.003***	0.013**	-0.005*	0.007*	0.587**
Strictly Exogenous Variables					
Some School (D)	0.015***	-0.190**	0.060	0.004***	-0.562*
Higher Secondary (D)	-0.033**	-0.076**	0.025**	0.009**	-0.002*
Some Voc. Train. (D)	-0.003***	0.184**	-0.015*	0.018**	0.067
University (D)	-0.054**	0.268***	-0.088**	0.049***	-0.012
Tech-Coll. (D)	-0.078*	0.156**	-0.187**	0.079***	0.537*
Job Exp. (Full T.)	0.003**	-0.143***	-0.016***	-0.005*	0.041
Age	-0.001**	0.064***	0.014***	-0.009***	-0.004
East German (D)	0.067***	-0.125**	-0.146***	-0.399***	0.248*
German (D)	-0.015**	0.170**	0.072**	0.013*	-0.368
Time Averages					
Some School (D)	-0.005***	0.265***	-0.009	0.006***	0.434*
Higher Secondary (D)	-0.036*	0.116***	-0.007*	0.015*	0.005
Some Voc. Train. (D)	0.007	-0.177***	0.025**	-0.007**	-0.005
University (D)	-0.061**	-0.153***	0.084***	-0.033***	0.261*
Tech-Coll. (D)	-0.012**	-0.118***	0.115*	-0.047**	-0.351*
Job Exp. (Full T.)	-0.005**	0.198*	0.012*	0.004*	-0.047*
Age	-0.001**	-0.072***	-0.026**	0.008***	-0.002
East German (D)	0.005***	0.054**	0.098***	0.168**	-0.319**
German (D)	-0.041***	-0.168*	-0.047***	-0.005	0.226
σ_a	0.346**	0.198***	0.451**	0.219*	0.571***
N					1,351

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Average Partial Effects for Headcount (Female) – Partial Likelihood Approach (1994 – 1996, Sampling Interval = 1 yr.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.048***	–	–	–	–
Children	0.011***	-0.159***	–	–	–
Partner in HH	-0.054***	-0.008**	0.001	–	–
Perceived Control (PC)	-0.010***	0.009***	-0.003***	0.009***	–
Lagged Variables					
Poverty St.	0.179***	0.017**	0.023***	-0.007***	-0.045*
Employment St.	0.019***	0.364***	-0.001	0.002	0.002
Children	0.009***	0.096***	0.878***	0.004***	-0.004
Partner in HH	0.064***	-0.012***	0.082***	0.219***	0.026*
Perceived Control (PC)	-0.006***	-0.005**	0.005***	-0.003**	0.793***
Strictly Exogenous Variables					
Some School (D)	-0.068***	-0.028**	0.191***	-0.035***	0.121*
Higher Secondary (D)	-0.061***	-0.016**	0.153***	-0.044***	0.196
Some Voc. Train. (D)	0.034***	0.543***	0.040***	0.058***	-0.208*
University (D)	0.016	0.495***	0.005**	0.008	0.192*
Tech-Coll. (D)	-0.068***	0.613***	0.137***	0.198***	-0.102
Job Exp. (Full T.)	-0.002	-0.105***	-0.002	-0.006**	0.001
Age	-0.012***	0.031***	0.004***	0.001	0.019*
East German (D)	-0.063***	0.179***	0.259***	0.035*	-0.087
German (D)	0.159***	0.119***	-0.216***	0.054***	-0.054
Time Averages					
Some School (D)	0.192***	0.098***	-0.103***	0.062***	-0.139
Higher Secondary (D)	0.136***	0.056***	-0.078***	0.069***	-0.135
Some Voc. Train. (D)	-0.044***	-0.389***	-0.010	-0.057***	0.271**
University (D)	-0.041***	-0.282***	0.028***	-0.010	-0.037
Tech-Coll. (D)	0.142***	-0.487***	-0.087***	-0.199***	0.250
Job Exp. (Full T.)	0.002	0.109***	0.001	0.005**	-0.001
Age	0.011***	-0.037***	-0.007***	-0.001	-0.022*
East German (D)	0.148***	-0.152***	-0.189***	-0.030	0.053
German (D)	-0.109***	-0.146***	0.119***	-0.040***	0.183
<i>N</i>	8,218	8,303	8,660	8,301	8,663

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Average Partial Effects for Headcount (Male) – Partial Likelihood Approach (1994 –1996, Sampling Interval = 1 yr.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.	PC
Endogenous Variables					
Employment St.	-0.042***	–	–	–	–
Children	0.054***	0.048***	–	–	–
Partner in HH	0.016***	0.006	0.243***	–	–
Perceived Control (PC)	-0.002***	0.006***	0.005***	0.003***	–
Lagged Variables					
Poverty St.	0.189***	0.019**	0.006*	-0.008***	-0.105**
Employment St.	0.040***	0.203***	0.024***	0.009***	0.028
Children	-0.027***	-0.012***	0.798***	0.004***	0.006
Partner in HH	-0.013***	0.032***	-0.045***	0.229***	0.068**
Perceived Control (PC)	-0.007***	0.003**	0.002***	-0.001**	0.378***
Strictly Exogenous Variables					
Some School (D)	-0.029***	-0.191***	0.016	0.002***	-0.507*
Higher Secondary (D)	0.002	-0.077***	0.021***	0.008***	-0.003
Some Voc. Train. (D)	-0.027***	0.181***	-0.012	0.017**	0.063
University (D)	0.029***	0.207***	-0.081***	0.045***	-0.013
Tech-Coll. (D)	-0.021	0.162***	-0.182***	0.085***	0.533**
Job Exp. (Full T.)	-0.014***	-0.143***	-0.019***	-0.003	0.047
Age	0.006***	0.069***	0.018***	-0.009***	-0.003
East German (D)	-0.041***	-0.101***	-0.152***	-0.352***	0.239
German (D)	-0.036***	0.151***	0.065***	0.011	-0.327
Time Averages					
Some School (D)	0.032***	0.207***	-0.008	0.009***	0.442*
Higher Secondary (D)	-0.008**	0.115***	-0.004	0.011***	0.012
Some Voc. Train. (D)	0.039***	-0.173***	0.026**	-0.010	-0.016
University (D)	-0.056***	-0.164***	0.083***	-0.030***	0.232
Tech-Coll. (D)	0.010	-0.125***	0.119***	-0.049***	-0.342*
Job Exp. (Full T.)	0.015***	0.147***	0.019***	0.002	-0.049*
Age	-0.006***	-0.075***	-0.022***	0.010***	-0.000
East German (D)	0.089***	0.055***	0.099***	0.143***	-0.303
German (D)	0.023**	-0.175***	-0.043***	-0.004	0.208
<i>N</i>	7,953	8,027	8,379	8,027	7,992

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Average Partial Effects for Headcount (Female) – Partial Likelihood Approach (1994 – 2010, Sampling Interval = 1 yr.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.
Endogenous Variables				
Employment St.	-0.062***	–	–	–
Children	-0.002*	-0.111***	–	–
Partner in HH	-0.056***	-0.015***	0.010***	–
Lagged Variables				
Poverty St.	0.368***	0.001	0.021***	-0.006***
Employment St.	0.029***	0.491***	-0.000	0.004***
Children	0.011***	0.061***	0.855***	-0.001**
Partner in HH	0.037***	-0.000	0.071***	0.273***
Strictly Exogenous Variables				
Some School (D)	0.013***	0.027***	0.042***	0.003**
Higher Secondary (D)	0.033***	0.027***	0.051***	0.013***
Some Voc. Train. (D)	-0.008***	0.094***	0.031***	0.010***
University (D)	-0.006**	0.193***	0.041**	0.032***
Tech-Coll. (D)	0.030***	0.113***	0.032***	0.036***
Job Exp. (Full T.)	-0.001***	-0.013***	0.002***	0.001***
Age	0.003***	0.000***	-0.002***	-0.001***
East German (D)	-0.014***	0.022***	-0.012***	-0.052***
German (D)	0.007**	-0.005*	0.002	0.007***
Time Averages				
Some School (D)	-0.018***	0.023***	-0.013***	0.003**
Higher Secondary (D)	-0.046***	0.022***	-0.019***	-0.009***
Some Voc. Train. (D)	-0.019***	-0.040***	-0.003**	0.001
University (D)	-0.054***	-0.075***	-0.014***	-0.022***
Tech-Coll. (D)	-0.066***	-0.052***	-0.016***	-0.030***
Job Exp. (Full T.)	-0.000	0.018***	-0.002***	-0.002***
Age	-0.003***	-0.006***	-0.000***	0.001***
East German (D)	0.058***	-0.029***	0.001	0.049***
German (D)	-0.038***	-0.023***	-0.026***	-0.019***
<i>N</i>	60,544	61,140	106,544	61,011

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Average Partial Effects for Headcount (Male) – Partial Likelihood Approach (1994 – 2010, Sampling Interval = 1 yr.)

Dependent Variables	Pov. St.	Emp. St.	Childn.	Partn.
Endogenous Variables				
Employment St.	-0.061***	–	–	–
Children	0.021***	0.025***	–	–
Partner in HH	-0.003**	0.022***	0.228***	–
Lagged Variables				
Poverty St.	0.339***	-0.003***	0.018***	-0.002***
Employment St.	0.035***	0.273***	0.027***	0.017***
Children	0.002**	-0.009***	0.801***	-0.003***
Partner in HH	-0.007***	0.025***	-0.038***	0.292***
Strictly Exogenous Variables				
Some School (D)	-0.009***	0.017***	0.006**	0.004***
Higher Secondary (D)	0.004***	0.037***	0.017***	0.012***
Some Voc. Train. (D)	0.006***	0.082***	0.015***	0.003***
University (D)	-0.030***	0.175***	-0.007**	0.017***
Tech-Coll. (D)	-0.026***	0.117***	-0.005*	0.017***
Job Exp. (Full T.)	-0.000	-0.007***	-0.003***	-0.002***
Age	0.003***	-0.000***	-0.001***	0.001***
East German (D)	0.003	-0.015***	0.011***	-0.023***
German (D)	0.002	-0.029***	0.035***	0.003*
Time Averages				
Some School (D)	0.014***	0.042***	0.002	0.001
Higher Secondary (D)	-0.001	0.018***	-0.009***	-0.002
Some Voc. Train. (D)	-0.016***	-0.027***	-0.005***	0.002**
University (D)	-0.028***	-0.073***	0.014***	-0.011***
Tech-Coll. (D)	-0.008***	-0.059***	0.010***	-0.005***
Job Exp. (Full T.)	-0.001***	0.011***	0.002***	0.003***
Age	-0.002***	-0.007***	-0.001***	-0.001***
East German (D)	0.023***	-0.017***	-0.026***	0.025***
German (D)	-0.029***	0.011***	-0.055***	-0.018***
<i>N</i>	58,089	58,617	100,681	58,469

Dummy variables are indicated by (D).

APEs for initial conditions from the Wooldridge-term are not reported. Standard errors underlying the reported significance are cluster robust on the individual level. The p-values for the APEs are approximated by the Delta-Method.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

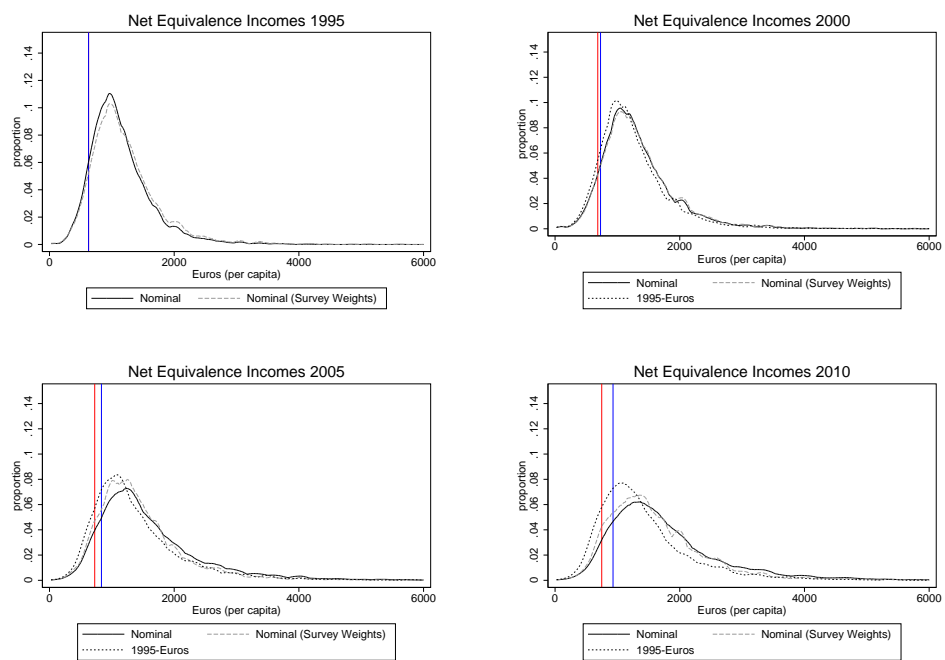


Figure 1: Development of Monthly Net Equivalence Incomes and the Poverty Line (Nominal = blue)

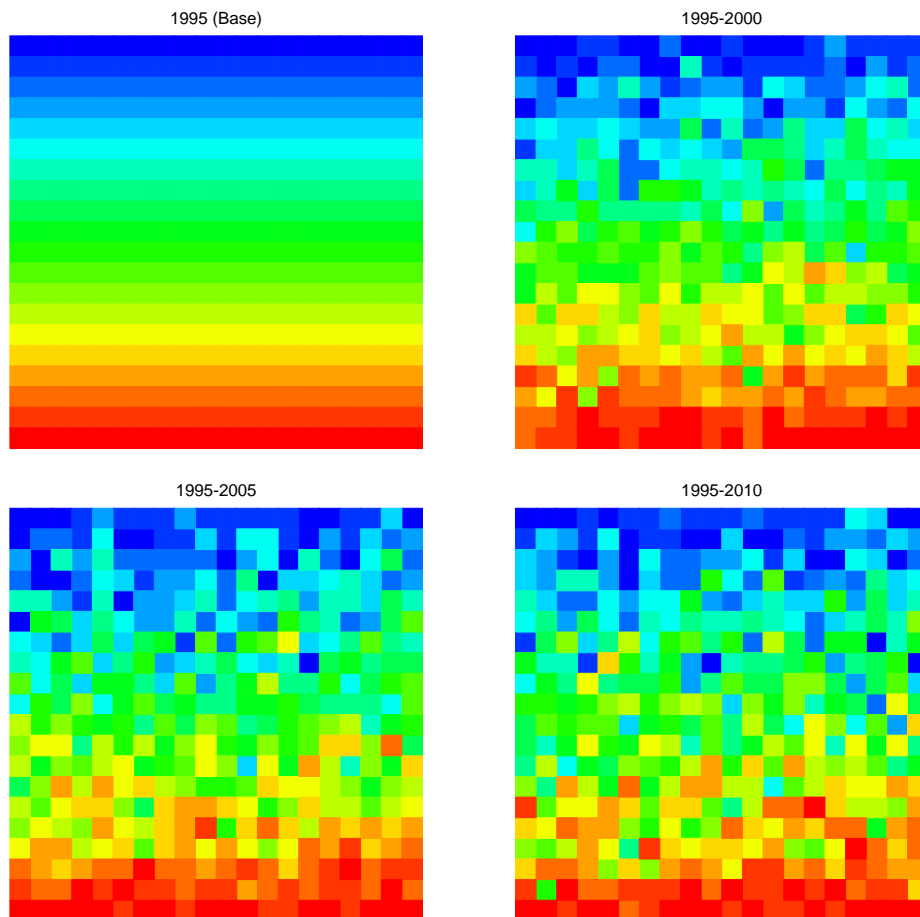


Figure 2: Mobility plots for net equivalence incomes with 1995 as a base period. The reference locations of the 1995-members for the 400 quantile increments in 2000, 2005, and 2010 are sorted in row-major order from left to right.

Appendix A: Personality traits as surveyed in the GSOEP

Perceived Control/ Locus of Control (LOC), as of 1999

Using the scale provided, indicate what your attitudes towards life and towards your own future are.

1. How my life goes depends on me (Internal LOC).
2. If a person is socially or politically active, he/she can have an effect on social conditions (Internal LOC).
3. One has to work hard in order to succeed (Internal LOC).
4. If I run up against difficulties in life, I often doubt my own abilities (reversed, Internal LOC).
5. Compared to other people, I have not achieved what I deserve (External LOC).
6. What a person achieves in life is above all a question of fate or luck (External LOC).
7. I frequently have the experience that other people have a controlling influence over my life (External LOC).
8. The opportunities that I have in life are determined by the social conditions (External LOC).
9. Inborn abilities are more important than any efforts one can make (External LOC).
10. I have little control over the things that happen in my life (External LOC).

Perceived Control, as of 1994

The following are various attitudes towards life and the future. Please indicate what most applies to you.

1. I determine what happens to me in life (Internal).
2. It is useless to make plans because they seldom work out (External).
3. My behavior determines my life (Internal).
4. No one can escape their fate, everything in life happens as it must happen (External).
5. If I get something I want then it's mostly due to luck (External).
6. Most plans I make are successful (Internal).
7. There is little sense in planing ahead because something unexpected always comes up (External).
8. Things always happen differently, one can't rely on anything (External).

Appendix B: Measurement of Perceived Control

Retrieving Unidimensional Item Combinations

To avoid erroneous estimates of latent control attitudes or, otherwise, not having to impose additional structure on the already complex estimation framework, it is particularly important to have an item combination available which is a one-dimensional representation of perceived control. To find the best item combination in terms of this objective, we apply an exploratory approach to a common factor model (see Anderson and Rubin, 1956). For each individual we obtain a measurement vector \mathbf{T}_i , which refers to a specific group of items that represent perceived control. However, there may potentially be more than one personality trait θ_i underlying the initial set of items \mathbf{T}_i . As such, the respective mean and covariance patterns for all $i \in N$ are

$$\begin{aligned}\mathbf{T} &= \mathbf{\Lambda}\boldsymbol{\theta} + \boldsymbol{\nu} \\ \mathbf{S} &= \mathbf{\Lambda}\boldsymbol{\Psi}\mathbf{\Lambda}' + \boldsymbol{\Theta},\end{aligned}\tag{B.1}$$

where $\boldsymbol{\Theta}$ is the covariance structure of the uniqueness $\boldsymbol{\nu}$. Furthermore, $\boldsymbol{\Psi} = \mathbf{I}$ implies that the common factors are a priori uncorrelated and have unit variance for the sake of identification.²⁵ It changes to a correlation matrix if the elements of S are normalized. The residuals $\boldsymbol{\nu}$ have mean zero and are uncorrelated with all $\boldsymbol{\theta}$ and among each other. Given these presumptions, the factor extraction is exclusively based on the observed matrix \mathbf{S} as it is already centralized around the corresponding observed means. In order to reduce potential statistical artifacts resulting from the categorical nature of the responses, it is common practice to take the underlying nature of \mathbf{S} into account. Such polychoric correlations for categorical item responses imply that the responses in \mathbf{T} are based on a latent

²⁵ Such normalization on either the factor loading or the corresponding factor variance are always required in factor models as the overall scale is otherwise unidentifiable (see Anderson and Rubin, 1956). In exploratory factor models, it is common to normalize the factor variance. The orthogonality assumption that prescribes $\boldsymbol{\Psi}$ to be diagonal can be replaced by other restrictions on $\mathbf{\Lambda}$, which are hard to reason in exploratory settings, however. As diagonalization leads to parsimony in terms of underlying factors and can be relaxed later on, it is common to proceed in this fashion. Moreover, if only few items are available, non-orthogonality between factor may prevent identification.

continuum $\tilde{\mathbf{T}}$. Hence, the correlations of the elements in $\tilde{\mathbf{T}}$ are to be estimated instead of those in \mathbf{T} . The diagonal elements of the resulting correlation matrix $\tilde{\mathbf{S}}$ are again unity, but half of the off-diagonal elements have to be evaluated. For these \tilde{S}_{ij} , suppose the underlying continuous \tilde{T}_i and \tilde{T}_j are

$$\begin{pmatrix} \tilde{T}_i \\ \tilde{T}_j \end{pmatrix} \sim \mathcal{BN} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{ij} \\ \rho_{ij} & 1 \end{pmatrix} \right),$$

where both, \tilde{T}_i and \tilde{T}_j , have mean zero and unit variance, implying

$$\Phi(\tilde{T}_i, \tilde{T}_j, \rho_{ij}) = \frac{1}{2\pi\sqrt{1-\rho_{ij}^2}} \int_{-\infty}^{\tilde{T}_i} \int_{-\infty}^{\tilde{T}_j} e^{\frac{1}{2(1-\rho_{ij}^2)}(\tilde{T}_i^2 - 2\rho_{ij}\tilde{T}_i\tilde{T}_j + \tilde{T}_j^2)} d\tilde{T}_i d\tilde{T}_j.$$

Given that for every two items (i, j) we have $k = 1 \dots K$ and $l = 1 \dots L$ response categories, where $K = L$ holds, one obtains the (i, j) -specific log-likelihood function

$$\ell_{ij} = \ln C + \sum_{k=1}^K \sum_{l=1}^K n_{kl} \ln \eta_{kl},$$

where η_{kl} is the cell probability of a response combination (k, l) which is observed n_{kl} times in the data. The contribution for η_{kl} is obtained from the double-difference of the cumulated density function defined above, where the intervals of the differences depend on unknown cutoff-points $\gamma_{i,k}$ and $\gamma_{j,l}$ ($k = 1 \dots K$, $l = 1 \dots L$). More specifically, we shall obtain

$$\eta_{kl} = [\Phi(\gamma_{i,k}, \gamma_{j,l}, \rho_{ij}) - \Phi(\gamma_{i,k-1}, \gamma_{j,l}, \rho_{ij})] - [\Phi(\gamma_{i,k}, \gamma_{j,l-1}, \rho_{ij}) - \Phi(\gamma_{i,k-1}, \gamma_{j,l-1}, \rho_{ij})],$$

which, when substituted into the above likelihood, provides full-information estimates of the respective ρ_{ij} .²⁶ This procedure is repeated for all triangular item

²⁶ There are also three-step procedures based on conditional likelihood estimates available, but combination-specific cutoff-estimates ought to perform better than row-specific first-stage estimates in most cases.

combinations in $\tilde{\mathbf{S}}$. Subsequently, the common factors that produce a high share of the common variance and, at the same time, a high number of retained items, can be extracted from the estimated $\hat{\tilde{\mathbf{S}}}$ by common methods. A convenient choice is the principal factor analysis with iterated communalities, for which the covariance structure in equation B.1 can be rewritten as follows:

$$\tilde{\mathbf{S}} - \Theta = \Lambda \Lambda'.$$

By definition, this step only affects the diagonal elements of $\tilde{\mathbf{S}}$, where the reduced values are called communalities. An initial estimate for the i -th communality \hat{h}_i^2 can be obtained from $1 - 1/r_{ii}$, where r_{ii} is the i -th diagonal element of the square matrix $\tilde{\mathbf{S}}^{-1}$ (see Mulaik, 2009). An estimate $\hat{\Lambda}$ of the factor loadings results from the factorization

$$\hat{\Lambda} \hat{\Lambda}' = \hat{\tilde{\mathbf{S}}} - \hat{\Theta} = \mathbf{C} \mathbf{D} \mathbf{C}' = \mathbf{C} \mathbf{D}^{1/2} \mathbf{D}^{1/2} \mathbf{C}',$$

which is the so-called spectral decomposition of the symmetric matrix $\hat{\tilde{\mathbf{S}}} - \hat{\Theta}$ with \mathbf{D} being the diagonal matrix of eigenvalues and \mathbf{C} being the matrix of the corresponding characteristic vectors.²⁷ Furthermore, as $\hat{\tilde{\mathbf{S}}} - \hat{\Theta}$ and \mathbf{D} are positive semi-definite, the factorization $\mathbf{D} = \mathbf{D}^{1/2} \mathbf{D}^{1/2}$ with $\mathbf{C} \mathbf{D}^{1/2} = \hat{\Lambda}$ applies. Since the factorized matrix is standardized, all factor loadings represent correlations and their item-specific sums are updated guesses of the communalities. Hence, the communalities in $\hat{\tilde{\mathbf{S}}} - \hat{\Theta}$ can be updated iteration-wise until they converge (see, e.g., Rencher, 2004). Sometimes the iterative nature leads to corner solutions. Such so-called Heywood cases (see, e.g., Thompson, 2004) are discarded and the respective second-best combinations are used instead. Following Costello and Osborne (2005), it is expedient to end up with a “clean” factor structure where the factor loadings associate as much items as possible with one major common factor explaining most of the variance.

Table A1 lays out the results for the finally selected item combinations. For

²⁷ Each column of \mathbf{C} forms a characteristic vector with orthonormalization such that $\mathbf{c}'_i \mathbf{c}_j = 0$ $\forall i \neq j$ and $\mathbf{c}'_i \mathbf{c}_j = 1 \forall i = j$.

waves 1999, 2005, and 2010, the same items are used and the resulting pattern is quite stable across waves. In order to account for potential gender differences, the common factor model is separately estimated for female and male sample members.

Table A1: Iterated Principle Factor Analysis for Polychoric Item Correlations of Perceive Control

		Common Variances			
		(female)		(male)	
		Eigenvalue	Proportion	Eigenvalue	Proportion
2010	1 st Principal Factor	1.671	0.896	1.805	0.876
	2 nd Principal Factor	0.143	0.077	0.131	0.064
2005	1 st Principal Factor	1.588	0.886	1.721	0.940
	2 nd Principal Factor	0.127	0.071	0.069	0.038
1999	1 st Principal Factor	1.699	0.939	1.761	0.959
	2 nd Principal Factor	0.074	0.041	0.052	0.029
1995	1 st Principal Factor	2.394	0.884	2.459	0.898
	2 nd Principal Factor	0.201	0.074	0.202	0.074
		Factor Structure			
		Factor 1	Factor 2	Factor 1	Factor 2
2010	Item 1	0.576	0.136	0.628	0.251
	Item 2	0.508	0.062	0.531	0.083
	Item 3	0.673	-0.106	0.669	-0.243
	Item 4	0.368	0.268	0.405	-0.014
	Item 5	0.702	-0.195	0.720	-0.047
2005	Item 1	0.526	0.101	0.591	0.058
	Item 2	0.527	0.079	0.523	0.143
	Item 3	0.647	-0.121	0.653	-0.113
	Item 4	0.363	0.264	0.398	0.131
	Item 5	0.696	-0.161	0.716	-0.123
1999	Item 1	0.616	0.056	0.649	-0.080
	Item 2	0.512	0.031	0.537	0.087
	Item 3	0.667	-0.027	0.658	0.061
	Item 4	0.357	0.212	0.366	0.140
	Item 5	0.697	-0.155	0.696	-0.123
1995	Item 1	0.626	0.017	0.657	-0.058
	Item 2	0.624	0.177	0.631	0.215
	Item 3	0.548	0.308	0.565	0.293
	Item 4	0.818	-0.252	0.824	-0.246
	Item 5	0.802	-0.104	0.795	-0.076

There are combinations with less than five retained items available that exhibit

slightly higher shares of common variance. Nonetheless, we opt for the five-item alternative as a higher number of measurement equations generally increases the quality of the factor scores derived later on. Apart from the share of common variance, the second objective of the item selection was to obtain homogeneous loadings on the first principal factor. Surprisingly, the item selection that complies most with both aims is the same for females and males. As mentioned in the data section, the 1994-to-1996 version of the GSOEP control inventory is a slightly different prequel of the later one. As the 1995 wave is predominantly used in the main model specifications, we jointly examine the factor structure and factor pattern of the 1995 and the 1999 item inventory.

If the item combinations for 1995 and 1999 are jointly evaluated, the share of the common variance is reduced by some 15 to 20 percentage points (Table A2). Instead, a second principal factor, which accounts for about 30 percent of the overall variance, occurs. The next common factor is again negligible. The loadings suggest that the association of all selected 1995 and 1999 items with respect to the principal projection axis is as intended. However, the second axis obviously implies a full reversal for the projection of both item blocks. Fortunately, this second dimension is almost orthogonal to the first factor, making the prequel version of the perceived control scale a still descent approximation to the later one.

Table A2: Factor Structure and Factor Pattern for the Association of the 1995 and 1999 Control-Perception Inventory

	Common Variances			
	(female)		(male)	
	Eigenvalue	Proportion	Eigenvalue	Proportion
1 st Principal Factor	2.791	0.693	2.982	0.713
2 nd Principal Factor	1.236	0.307	1.199	0.287
3 rd Principal Factor	0.110	0.027	0.142	0.034
	Factor Structure			
	Factor 1	Factor 2	Factor 1	Factor 2
Item 1 (1995)	0.584	-0.193	0.619	-0.221
Item 2 (1995)	0.538	-0.319	0.535	-0.324
Item 3 (1995)	0.509	-0.160	0.542	-0.121
Item 4 (1995)	0.721	-0.343	0.737	-0.340
Item 5 (1995)	0.741	-0.288	0.730	-0.305
Item 1 (1999)	0.412	0.451	0.471	0.447
Item 2 (1999)	0.425	0.318	0.435	0.316
Item 3 (1999)	0.396	0.578	0.399	0.562
Item 4 (1999)	0.268	0.202	0.326	0.200
Item 5 (1999)	0.504	0.440	0.516	0.408
	Factor Pattern (after Rotation)			
Item 1 (1995)	0.553	-0.270	0.589	-0.292
Item 2 (1995)	0.490	-0.387	0.494	-0.384
Item 3 (1995)	0.483	-0.226	0.524	-0.183
Item 4 (1995)	0.669	-0.436	0.692	-0.423
Item 5 (1995)	0.696	-0.384	0.689	-0.388
Item 1 (1999)	0.469	0.391	0.520	0.389
Item 2 (1999)	0.464	0.259	0.469	0.263
Item 3 (1999)	0.470	0.520	0.462	0.512
Item 4 (1999)	0.292	0.165	0.347	0.161
Item 5 (1999)	0.558	0.368	0.561	0.345

Factor structure refers to the factor loadings under factor orthogonality that can be seen as correlation coefficients. After rotation, the reported coefficients are only interpretable as *factor pattern/weights* and do not represent correlations any longer.

Estimating Latent Control-Attitudes

Given the unidimensionality of the factors underlying our item selection, the next step is to estimate the latent traits for each individual. We use a nonparametric item response model introduced by Spady (2007). Given that there is a positive relation between latent control attributes and the corresponding responses on a K -

point Likert-scale, Item Response Models are used to estimate the mappings from θ into the probability range $[0, 1]$ for all $k = 1, \dots, K$ responses. The probability of response K approaches zero for low θ and one for very high θ . Conversely, the probability of giving response $k = 1$ is 1 for very low levels of θ and 0 for high levels. Responses $k = 2$ to $k = K - 1$ have bell-shaped probabilities with their location shifting from left to right as k increases. For the ease of identification (see Spady, 2006, for exact conditions) however, it is meaningful to model cumulated response probabilities $P(r \leq k|\theta)$ instead of $P(r = k|\theta)$. As $P(r \leq K|\theta) = 1$ always holds, $K - 1$ cumulated response curves that are monotonically decreasing in θ and non-intersecting have to be estimated. Non-intersection is bound to arise by setting up the cumulated response probabilities as follows:

$$\begin{aligned}
P(r \leq K - 1|\theta) &= 1 - G(K - 1|u) \\
P(r \leq K - 2|\theta) &= [1 - G(K - 2|u)] P(r \leq K - 1|\theta) \\
&\vdots \\
P(r \leq 1|\theta) &= [1 - G(1|u)] P(r \leq 2|\theta)
\end{aligned} \tag{B.2}$$

Since $P(r \leq K - 1|\theta)$ is decreasing in θ , $G(K - 1|u)$ is increasing in it via some monotone mapping u . Spady (2007) establishes general conditions under which polynomial series are a flexible way to approximate the respective $G(k|u)$ for $k = 1, \dots, K$. This approach is what renders our estimation strategy a non-parametric Item Response Model (see Härdle, 1995, Chen, 2007, for an overview on nonparametric estimators based on orthogonal series). More explicitly, we use shifted Legendre polynomials of the third degree (see Judd, 1998) to approximate $G(k|u)$ using an *exponential tilting* factor as in Barron and Sheu (1991).

$$G(u) = \frac{\int_0^\theta e^{t_1\gamma_1(u)+t_2\gamma_2(u)+t_3\gamma_3(u)} du}{\int_0^1 e^{t_1\gamma_1(u)+t_2\gamma_2(u)+t_3\gamma_3(u)} du}, \tag{B.3}$$

where the transform $u = \Phi(\theta)$ is used to match the support of θ with the domain $[0, 1]$ of the polynomial basis, and t_1 to t_3 are the parameters to be estimated for each θ . The bound θ is a placeholder for the respective $u = \Phi(\theta)$. Hence, every

$P(r = k|\theta)$ can simply be expressed as differences of the respective cumulated response curves as modeled in equation B.2. The resulting likelihood contribution for the i th individual is

$$\begin{aligned} p_i(r_{i1}, r_{i2}, \dots, r_{im}) &= \int p_{ip}(r_{i1}, r_{i2}, \dots, r_{im}|\theta_i) f(\theta_i) d\theta_i \\ &= \int p_{i1}(r_{i1}|\theta_i) p_{i2}(r_{i2}|\theta_i) \dots p_{im}(r_{im}|\theta_i) f(\theta_i) d\theta_i. \end{aligned} \tag{B.4}$$

The second expression requires local independence, which states that all individual characteristics that may influence the response probabilities are conveyed in θ . Hence, response probabilities across items $j = 1, \dots, m$ are independent conditional on θ . In order to obtain a Likelihood expression unconditional on unobserved traits, θ_i can easily be integrated out by assuming $f(\theta)$ to be $\mathcal{N}(0, 1)$ distributed.

Given the estimated polynomial coefficients obtained from equation B.4, we can predict an individual's θ_i by an *Empirical Modal Bayes* approach (see SkronDAL and Rabe-Hesketh, 2004). Each response pattern uniquely determines θ_i by finding the mode of the implied empirical posterior.

$$f(\theta_i|r) = \frac{f(\theta, r)}{p(r)} = \frac{p_1(r_1|\theta) \dots p_m(r_m|\theta) f(\theta)}{\int p_1(r_1|\theta) p_2(r_2|\theta) \dots p_m(r_m|\theta) f(\theta) d\theta} \tag{B.5}$$