I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14317

Structural Models for Policy-Making: Coping with Parametric Uncertainty

Philipp Eisenhauer
Janoś Gabler
Lena Janys

APRIL 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14317

Structural Models for Policy-Making: Coping with Parametric Uncertainty

Philipp Eisenhauer
University of Bonn

Janoś Gabler
University of Bonn and IZA

Lena Janys
University of Bonn, HCM and IZA

APRIL 2021

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Structural Models for Policy-Making: Coping with Parametric Uncertainty*

The ex-ante evaluation of policies using structural econometric models is based on estimated parameters as a stand-in for the truth. This practice ignores uncertainty in the counterfactual policy predictions of the model. We develop a generic approach that deals with parametric uncertainty using uncertainty sets and frames model-informed policymaking as a decision problem under uncertainty. The seminal human capital investment model by Keane and Wolpin (1997) provides us with a well-known, influential, and empirically-grounded test case. We document considerable uncertainty in their policy predictions and highlight the resulting policy recommendations from using different formal rules on decision-making under uncertainty.

JEL Classification: C44, C54, D81

Keywords: statistical decision rules, structural estimation, uncertainty quantification

Corresponding author:
Philipp Eisenhauer
University of Bonn
Institute for Applied Microeconomics
Adenauerallee 24-42
53113 Bonn
Germany
E-mail: peisenha@uni-bonn.de

* Philipp Eisenhauer and Lena Janys are both funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1- 390838866 and the TRA Modelling (University of Bonn) as part of the Excellence Strategy of the federal and state governments. Janoś Gabler is grateful for financial support by the German Research Foundation (DFG) through CRC-TR 224 (Project C01) and funding by IZA Institute of Labor Economics. Lena Janys is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2047/1 - 390685813. Philipp Eisenhauer was funded by a postdoctoral fellowship by the AXA Research Fund. We thank Tim Mensinger for his help in the early stages of the project. We thank Max Blesch, Joachim Freyberger, Annica Gehlen, Daniel Harenberg, Ken Judd, Gregor Reich, Jörg Stoye, and Rafael Suchy for numerous helpful discussions. We thank Michael Keane and Kenneth Wolpin for providing the dataset used in our analysis. We thank Annica Gehlen for her outstanding research assistance. We are grateful to the Social Sciences Computing Service (SSCS) at the University of Chicago for the permission to use their computational resources.

Contents

1. Introduction

Economists use highly parameterized structural models to investigate economic mechanisms, predict the impact of proposed policies, and inform optimal policy-making (Wolpin, 2013). These models represent deep structural relationships of theoretical economic models invariant to policy changes (Hood and Koopmans, 1953). The sources of uncertainty in such an analysis are ubiquitous (Saltelli et al., 2020). For example, all models are misspecified, there are numerical approximation errors in their implementation, and model parameters are uncertain. Therefore, a proper accounting of uncertainty is a prerequisite for using computational models for decision-making in most disciplines (National Research Council, 2012; SAPEA, 2019).

Our focus is on parametric uncertainty in structural econometric models that are estimated on observed data. Economists often ignore parametric uncertainty and conduct an as-if analysis where the point estimates serve as a stand-in for the true model parameters. We then continue to study the implications of our models at the point estimates (Adda et al., 2017; Blundell et al., 2016; Eckstein et al., 2019; Eisenhauer et al., 2015) and rank competing policy proposals based on the point predictions only (Blundell and Shephard, 2012; Cunha et al., 2010; Gayle and Shephard, 2019; Todd and Wolpin, 2006). In fact, in their handbook article, Keane et al. (2011) state that they are unaware of any applied work that reported the distribution of policy predictions due to parametric uncertainty. To the best of our knowledge, this statement remains true more than a decade later. Consequently, economists run the risk of accepting fragile findings as facts, ignoring the trade-off between model complexity and prediction uncertainty, and not framing policy advice as a decision problem under uncertainty.

In this paper, we develop an approach that copes with parametric uncertainty and embeds model-informed policy-making in a decision-theoretic framework. We follow Manski (2021)'s suggestion and, instead of using the parameter estimates as-if they were true, incorporate uncertainty in the analysis by treating the estimated confidence set as-if it is correct. We use the confidence set to construct an uncertainty set that is anchored in empirical estimates, statistically meaningful, and computationally tractable (Ben-Tal et al., 2013). Instead of just focusing on the point estimates, we evaluate counterfactual policies based on all parametrizations within the uncertainty set.

We rely on statistical decision theory (Manski, 2013) to deal with the uncertainty in counterfactual predictions. This approach promotes a well-reasoned and transparent policy process. Before a decision, it clarifies trade-offs between choices (Gilboa et al., 2018). Afterward, decision-theoretic principles allow constituents to scrutinize the coherence of choices (Gilboa, 2020), ease the ex-post justification (Berger et al., 2020), and facilitate the communication of uncertainty (Manski, 2019).

As an example of our generic approach, we analyze the seminal human capital investment model by Keane and Wolpin (1997) as a well-known, empirically grounded, and computationally demanding test case. We follow the authors and estimate the model on the National Longitudinal Survey of Youth 1979 (NLSY79) (Bureau of Labor Statistics, 2019) using the original dataset and reproduce all core results. We revisit their predictions for the impact of a tuition subsidy on completed years of schooling. The economics of the model imply that the nonlinear mapping between the model parameters and predictions is truncated at zero. We thus use the Confidence Set (CS) bootstrap (Woutersen and Ham, 2019) to estimate the confidence set for the counterfactuals. We document considerable uncertainty in the policy predictions and highlight the resulting policy recommendations from different formal rules on decision-making under uncertainty.

Our work extends existing research exploring the sensitivity of implications and predictions to parametric uncertainty in macroeconomics and climate economics. For example, Harenberg et al. (2019) study uncertainty propagation and sensitivity analysis for a standard real business cycle model. Cai and Lontzek (2019) examine how uncertainties and risks in economic and climate systems affect the social cost of carbon. However, neither of them estimates their model on data. Instead they rely on expert judgments to inform the degree of parametric uncertainty. They do not investigate the consequences of uncertainty for policy decisions in a decision-theoretic framework.

Our work complements a burgeoning literature on the sensitivity analysis of policy predictions in light of model or moment misspecification. For example, Andrews et al. (2017) and Andrews et al. (2020) treat the model specification as given and then analyze the sensitivity of the parameter estimates to misspecification of the moments used for estimation. Christensen and Connault (2019) study global sensitivity of the model predictions to misspecification of the distribution of unobservables. Jørgensen (2021) provides a local measure for the sensitivity of counterfactuals to model parameters that are fixed before the estimation of the model.[1] This literature does not embed the counterfactual predictions in a decision-thoretic setting.

We structure the remainder of this paper as follows. We first describe the decision-theoretic frame-

[1] See for other examples Armstrong and Kolesár (2021), Bonhomme and Weidner (2020), Bugni and Ura (2019), and Mukhin (2018).

work for making model-informed decisions under parametric uncertainty in Section 2 using an illustrative example. We then summarize the empirical setting of Keane and Wolpin (1997) in Section 3, before Section 4 presents our results before. We finish in Section 5 with a brief conclusion and outlook.

2. Structural models for policy-making

We now outline the basic idea behind uncertainty propagation, provide more details on the current practice to use the estimated parameters as a plug-in replacement for the truth, and explore its limitations. We then discuss how to use the estimated confidence sets to construct uncertainty sets. The use of uncertainty sets allows us to cope with uncertain policy predictions in a proper decision-theoretic framework.

At a high level, a structural econometric model provides a mapping $\mathcal{M}(\boldsymbol{\theta})$ between the l model parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and a quantity y that is of interest to policy-makers.

$$\mathbb{R}^l \supset \boldsymbol{\Theta} \ni \boldsymbol{\theta} \mapsto \mathcal{M}(\boldsymbol{\theta}) = y$$

A policy $g \in \mathcal{G}$ changes the mapping to $\mathcal{M}_g(\boldsymbol{\theta})$ and produces a counterfactual y_g.

Estimation of a baseline model – describing the status-quo – on observed data allows researchers to learn about the true parameters. Frequentist estimation procedures such as maximum likelihood estimation or the method of simulated moments produce a point estimate $\hat{\boldsymbol{\theta}}$. However, uncertainty about the true parameters remains.

Previewing our empirical analysis of Keane and Wolpin (1997), our $\mathcal{M}$ is provided by a dynamic model of human capital accumulation, which we estimate on observed schooling and labor market decisions using simulated maximum likelihood estimation. The policy g is the implementation of a college tuition subsidy, and the counterfactual is the level of completed schooling in the population. Example parameters that drive the economics of the model are time preferences of individuals, the return to schooling, and the transferability of work experience across occupations.

We use an illustrative example throughout this section to highlight our key points. We consider two policies $g \in \{1, 2\}$ that result in two different mappings $(\mathcal{M}_1, \mathcal{M}_2)$ of the same scalar θ to a counterfactual y_g. The point estimate $\hat{\theta}$ is determined by estimating a baseline model on an observed dataset. We denote the probability density function of its sampling distribution by $f_{\hat{\theta}}$.

Under the first policy, the counterfactual is an increasing nonlinear function of θ. In the case of the second policy, the relationship is decreasing and linear.

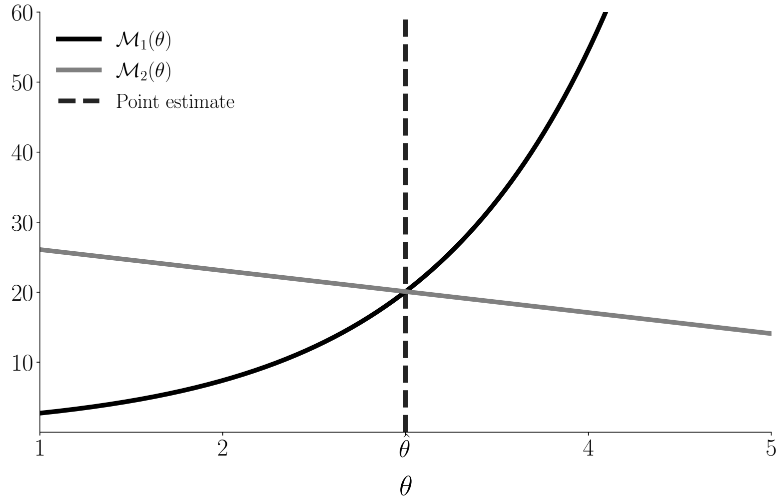


Figure 1.: Model comparison

Notes: We parameterize the two models as $y_1 = \exp\theta$ and $y_2 = 29.08 - 3\,\theta$.

Figure 1 traces the counterfactual from both models over a range of the parameter. At the point estimate, both models yield the same value for the counterfactual. For higher values, the first policy is preferred, while the opposite is true for lower values. So, once we take uncertainty about the true parameter into account, deciding which policy to adopt is not straightforward.

2.1. Uncertainty sets

Manski (2021) suggests acknowledging parametric uncertainty by working with estimated confidence sets instead of point estimates. A confidence set $\mathbf{\Theta}(\alpha) \subset \mathbf{\Theta}$ covers the true parameters, from an ex-ante point of view, with a predetermined coverage probability of $(1-\alpha)$. Going forward, instead of using the estimated parameter values as-if they were true, we analyze policy decisions using the estimated confidence set for the parameters $\hat{\mathbf{\Theta}}(\alpha)$ and the counterfactual $\hat{\mathbf{\Theta}}_{y_g}(\alpha)$ as-if it is correct.

Based on the estimated confidence sets, we construct so called uncertainty sets for the parameters $\mathcal{U}(\alpha)$ and the prediction $\mathcal{U}_{y_g}(\alpha)$ by only considering parameterizations that we cannot reject based on a hypothesis test with confidence level $1-\alpha$. This approach ensures the tractability of our decision-theoretic analysis as the uncertainty set of the parameters is much smaller than the whole parameter space of the model. We adopt this procedure from the literature on data-driven robust optimization in operations research (Ben-Tal et al., 2013; Bertsimas et al., 2018).

2.2. Statistical decision theory

In our setting, a policy-maker relies on a structural model with an uncertain parametrization to map alternative policies to counterfactual predictions. In most cases, the preferred policy depends on the model's uncertain parameters, and we draw on statistical decision theory to organize the decision-making process (Gilboa, 2009; Marinacci, 2015).

Returning to our example, we rank the two policies according to alternative statistical decision rules using an uncertainty set derived from a confidence set with a 90% coverage probability. In what follows, we postulate a linear utility function $U(y_g)$ describing the policy maker's preferences for simplicity.[2]

Figure 2 shows the implied sampling distribution of the predictions for the two alternative policies and the corresponding uncertainty sets $\mathcal{U}_{y_g}(0.1)$. The mapping $\mathcal{M}_1$ is highly nonlinear, while the mapping $\mathcal{M}_2$ is linear. When evaluated at the point estimate, the counterfactual is the same under both policies, so a policy-maker is indifferent. However, the spread of the uncertainty set differs considerably.

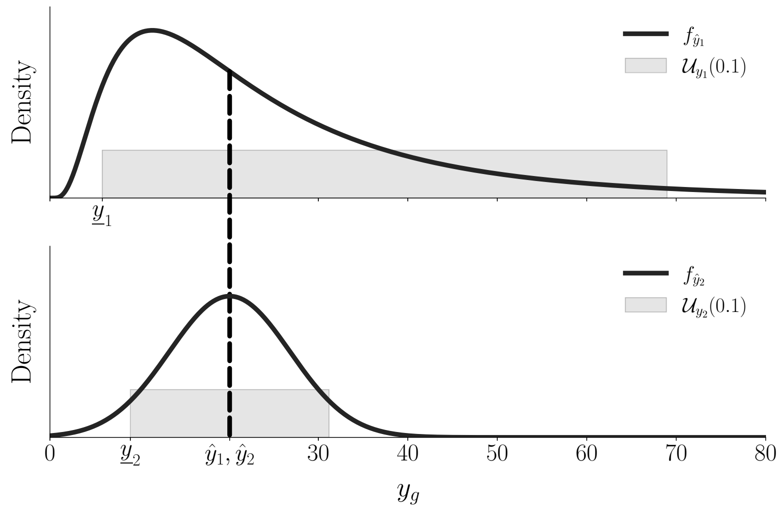


Figure 2.: Comparing policy predictions

Decision theory proposes a variety of different rules for reasonable decisions in this setting. We ex-

[2]We assume that the sampling distribution of the point estimate is normal with a mean of three and a standard deviation of three-fourth. We can derive the uncertainty sets directly and simply consider realizations of $\theta \in [1.76, 4.23]$.

plore the following four: (1) as-if optimization, (2) maximin criterion, (3) minimax regret rule, and (4) subjective Bayes.

As-if optimization describes the predominant practice. The estimation of the model produces point estimates that serve as a plug-in for the true parameters. The decision maximizes the utility at the point estimate. More formally,

$$g^* = \arg\max_{g \in \mathcal{G}} U(M_g(\hat{\boldsymbol{\theta}})).$$

Given our example, an as-if policy-maker is indifferent between the two policies. Both result in the same counterfactual at the point estimates as indicated by the dashed line in Figure 2.

The maximin criterion and minimax regret rule are two common alternatives favoring actions that work uniformly well over all possible parameters in the uncertainty set. This approach departs from as-if optimization, which only considers a policy's performance at a single point in the uncertainty set. The maximin decision (Wald, 1950; Gilboa and Schmeidler, 1989) is determined by computing the minimum utility for each policy within the uncertainty set and choosing the one with the highest worst-case outcome. Stated concisely,

$$g^* = \arg\max_{g \in \mathcal{G}} \min_{\boldsymbol{\theta} \in \mathcal{U}(\alpha)} U(M_g(\boldsymbol{\theta})).$$

In our example, and again returning to Figure 2, a maximin policy-maker prefers g_2 as its worst-case outcome within the uncertainty set $\underline{y}_2$ is better than under the alternative policy g_1.

The minimax regret rule (Niehans, 1948; Manski, 2004) computes the maximum regret for each policy over the whole uncertainty set and chooses the policy that minimizes the maximum regret. The regret of choosing a policy g for a given parameterization of the model is the difference in utility between the maximum possible utility from adopting $\tilde{g} \in \mathcal{G}$ and the actual utility received. The decision maximizes:

$$g^* = \arg\min_{g \in \mathcal{G}} \max_{\boldsymbol{\theta} \in \mathcal{U}(\alpha)} \underbrace{\left[\max_{\tilde{g} \in \mathcal{G}} U(M_{\tilde{g}}(\boldsymbol{\theta})) - U(M_g(\boldsymbol{\theta}))\right]}_{\text{regret}}.$$

Figure 3 compares our two policy examples over the uncertainty sets. A policy-maker adopting policy g_1 regrets his choice for small values of the model parameter, while the opposite is true for larger values. The regret of each policy is maximized at the boundaries of the uncertainty set.

Maximum regret is minimized when a policy-maker chooses g_1. It corresponds to the difference in the counterfactual at the lower boundary of the uncertainty set instead of the larger difference at its upper bound. This outcome contradicts the maximin decision where policy g_2 is preferred.

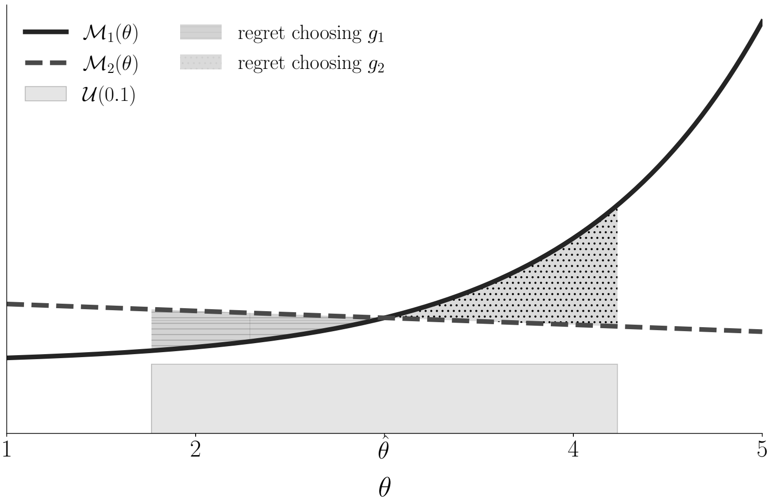


Figure 3.: Comparing policy regret

All decision rules presented so far focus on a single point in the uncertainty set as the policy's relevant performance measure. Bayesian approaches aggregate a policy's performance over the complete uncertainty set.

Maximization of the subjective expected utility (Savage, 1954) requires the policy-maker to place a subjective probability distribution $f_{\boldsymbol{\theta}}$ over the parameters in the uncertainty set. Then a policy-maker selects the alternative with the highest expected subjective utility. More formally,

$$g^* = \arg\max_{g \in \mathcal{G}} \int_{\mathcal{U}(\alpha)} U(M_g(\boldsymbol{\theta}))\, \mathrm{d}f_{\boldsymbol{\theta}}.$$

Applying a uniform distribution to our example, a policy-maker chooses g_1 as it performs well for high values of $\boldsymbol{\theta}$ and still reasonably well for low values.

3. Eckstein-Keane-Wolpin models

We now apply our framework to an Eckstein-Keane-Wolpin (EKW) model (Aguirregabiria and Mira, 2010) to investigate the prevalence and consequences of parametric uncertainty in an empirically grounded and computationally demanding setting. EKW models are often used by labor economists to learn about human capital investment and consumption-saving decisions and predict the impact of proposed reforms to education policy and welfare programs (Keane et al., 2011; Low and Meghir,

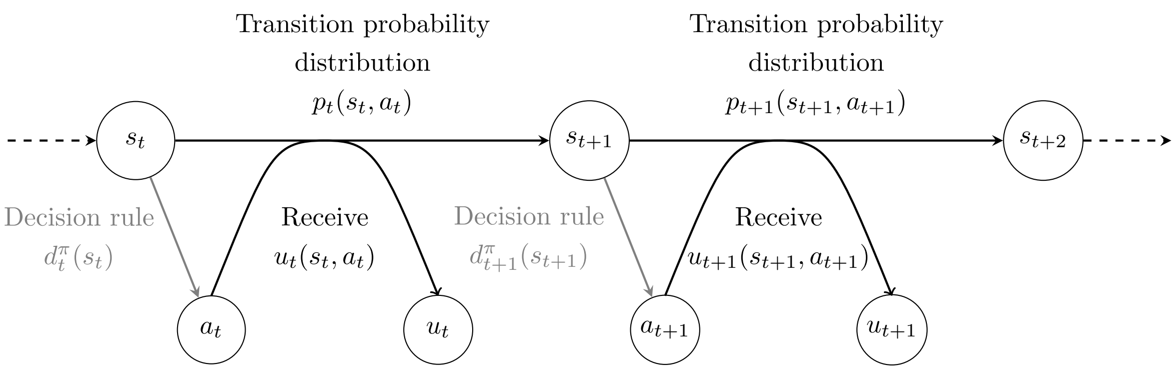


Figure 4.: Timing of events

2017; Blundell, 2017). We start by presenting the general structure of this class of models and their solution approach. We then turn to the customized version used by Keane and Wolpin (1997) to study the career decisions of young men. We outline their model's basic setup, provide some descriptive statistics of the empirical data used in our estimation, and then discuss the core findings.

3.1. Model structure

EKW models describe sequential decision-making under uncertainty (Gilboa, 2009; Machina and Viscusi, 2014). At time $t = 1, \ldots, T$ each individual observes the state of their choice environment $s_t \in S$ and chooses an action a_t from the set of admissible actions $\mathcal{A}$. The decision has two consequences: an individual receives an immediate utility $u_t(s_t, a_t)$ and their environment evolves to a new state s_{t+1}. The transition from s_t to s_{t+1} is affected by the action but remains uncertain. Individuals are forward-looking. Thus they do not simply choose the alternative with the highest immediate utility. Instead, they take the future consequences of their current action into account.

A policy $\pi = (d_1^\pi, \ldots, d_T^\pi)$ provides the individual with instructions for choosing an action in any possible future state. It is a sequence of decision rules d_t^π that specify the action $d_t^\pi(s_t) \in \mathcal{A}$ at a particular time t for any possible state s_t under π. The implementation of a policy generates a sequence of utilities that depends on the objective transition probability distribution $p_t(s_t, a_t)$ for the evolution of state s_t to s_{t+1} induced by the model.

Figure 4 depicts the timing of events for two generic time periods. At the beginning of period t, an individual fully learns about each alternative's immediate utility, chooses one of them, and receives its immediate utility. Then the state evolves from s_t to s_{t+1} and the process is repeated in $t + 1$.

Individuals make their decisions facing uncertainty about the future and seek to maximize their expected total discounted utilities over all decision periods given all available information. They have rational expectations (Muth, 1961), so their subjective beliefs about the future agree with the objective probabilities for all possible future events provided by the model. Immediate utilities are separable between periods (Kahneman et al., 1997), and a discount factor δ parameterizes a taste for immediate over future utilities (Samuelson, 1937).

Equation (1) provides the formal representation of the individual's objective. Given an initial state s_1, they implement a policy π that maximizes the expected total discounted utilities over all decision periods given the information available at the time.

$$\max_{\pi \in \Pi} \mathrm{E}^{\pi}_{s_1}\left[\sum_{t=1}^{T} \delta^{t-1} u_t(s_t, d^{\pi}_t(s_t))\right] \tag{1}$$

EKW models are set up as a standard Markov decision process (MDP) (Puterman, 1994; White, 1993; Rust, 1994) that can be solved by a simple backward induction procedure. In the final period T, there is no future to consider, and the optimal action is choosing the alternative with the highest immediate utility in each state. With the decision rule for the final period, we can determine all other optimal decisions recursively. We use our group's open-source research code `respy` (Gabler and Raabe, 2020) that allows for the flexible specification, simulation, and estimation of EKW models. Detailed documentation of the software and its numerical components is available at `http://respy.readthedocs.io`.

3.2. The career decisions of young men

Keane and Wolpin (1997) specialize the model above to explore the career decisions of young men regarding their schooling, work, and occupational choices using the National Longitudinal Survey of Youth 1979 (NLSY79) (Bureau of Labor Statistics, 2019) for the estimation of the model. We restrict ourselves to a basic summary of their setup. More detailed documentation of the model specification and the observed dataset is available in our Appendix.

Keane and Wolpin (1997) follow individuals over their working life from young adulthood at age 16 to retirement at age 65. The decision period $t = 16, \dots, 65$ is a school year. Figure 5 illustrates the initial decision problem as individuals decide $a \in \mathcal{A}$ whether to work in a blue-collar or white-collar occupation ($a = 1, 2$), to serve in the military ($a = 3$), to attend school ($a = 4$), or to stay at home ($a = 5$).

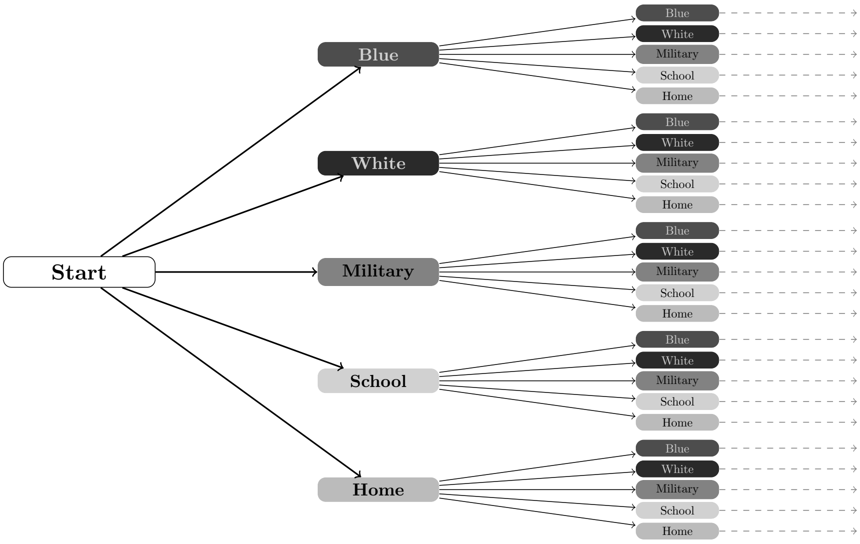


Figure 5.: Decision tree

Individuals are already heterogeneous when entering the model. They differ with respect to their level of initial schooling h_{16} and have one of four different $\mathcal{J} = \{1, \dots, 4\}$ alternative-specific skill endowment types $\boldsymbol{e} = (e_{j,a})_{\mathcal{J} \times \mathcal{A}}$.

The immediate utility $u_a(\cdot)$ of each alternative consists of a non-pecuniary utility $\zeta_a(\cdot)$ and, at least for the working alternatives, an additional wage component $w_a(\cdot)$. Both depend on the level of human capital as measured by their alternative-specific skill endowment $\boldsymbol{e}$, years of completed schooling h_t, and occupation-specific work experience $\boldsymbol{k_t} = (k_{a,t})_{a \in \{1,2,3\}}$. The immediate utilities are influenced by last-period choices a_{t-1} and alternative-specific productivity shocks $\boldsymbol{\epsilon_t} = (\epsilon_{a,t})_{a \in \mathcal{A}}$ as well. Their general form is given by:

$$u_a(\cdot) = \begin{cases} \zeta_a(\boldsymbol{k_t}, h_t, t, a_{t-1}) + w_a(\boldsymbol{k_t}, h_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if}\, a \in \{1, 2, 3\} \\ \zeta_a(\boldsymbol{k_t}, h_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if}\, a \in \{4, 5\}. \end{cases}$$

Work experience $\boldsymbol{k_t}$ and years of completed schooling h_t evolve deterministically. There is no uncertainty about grade completion (Altonji, 1993) and no part-time enrollment. Schooling is defined as time spent in school and not by formal credentials acquired. Once individuals reach a certain amount of schooling, they acquire a degree.

$$\begin{aligned} k_{a,t+1} &= k_{a,t} + \mathbf{I}[a_t = a] \qquad \text{if}\, a \in \{1, 2, 3\} \\ h_{t+1} &= h_t \;\; + \mathbf{I}[a_t = 4] \end{aligned}$$

The productivity shocks $\boldsymbol{\epsilon_t}$ are uncorrelated across time and follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. Given the structure of the utility functions and the distribution of the shocks, the state at time t is $s_t = \{\boldsymbol{k_t}, h_t, t, a_{t-1}, \boldsymbol{e}, \boldsymbol{\epsilon_t}\}$.

Skill endowments $\boldsymbol{e}$ and initial schooling h_{16} are the only sources of persistent heterogeneity in the model. All remaining differences in life-cycle decisions result from different transitory shocks $\boldsymbol{\epsilon_t}$ over time.

Theoretical and empirical research from specialized disciplines within economics informs the specification of each $u_a(\cdot)$, and we discuss the exact functional form of the non-pecuniary utility from schooling as an example in Equation (6). Further details on the specification of the utility functions is available in our Appendix.

$$\zeta_4(s_t) = \underbrace{e_{j,4}}_{\text{type}} + \underbrace{\beta_{tc_1} \cdot \mathbf{I}[h_t \geq 12] + \beta_{tc_2} \cdot \mathbf{I}[h_t \geq 16]}_{\text{tuition costs}} + \underbrace{\gamma_{4,4} \cdot t + \gamma_{4,5} \cdot \mathbf{I}[t < 18]}_{\text{time trend}} \tag{2}$$

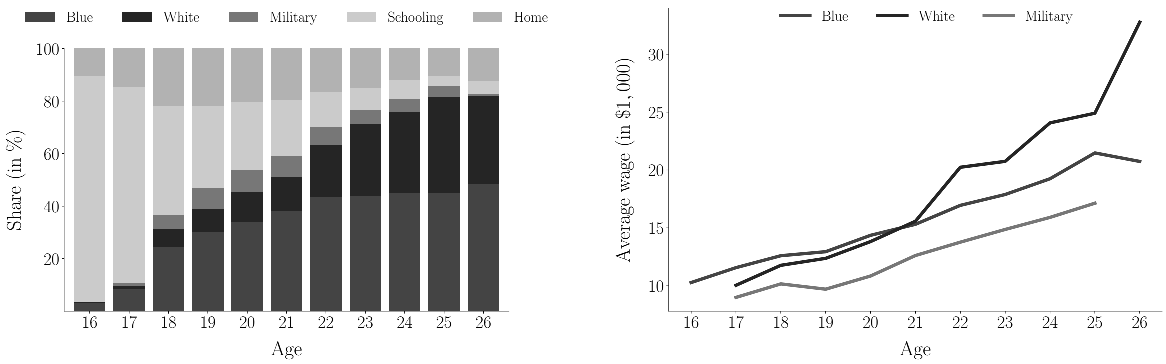
$$+ \underbrace{\beta_{rc_1} \cdot \mathbf{I}[a_{t-1} \neq 4, h_t < 12] + \beta_{rc_2} \cdot \mathbf{I}[a_{t-1} \neq 4, h_t \geq 12]}_{\text{re-enrollment cost}} + \ldots + \epsilon_{4,t}$$

There is a direct cost of attending school such as tuition for continuing education after high school β_{tc_1} and college β_{tc_2}. The decision to leave school is reversible, but entails re-enrollment costs that differ by schooling category $(\beta_{rc_1}, \beta_{rc_2})$.

We analyze the original dataset used by Keane and Wolpin (1997). We only provide a brief description and relegate further details to our Appendix. The authors construct their sample based on the NLSY79, a nationally representative sample of young men and women living in the United States in 1979 and born between 1957 and 1964. Individuals were followed from 1979 onwards and repeatedly interviewed about their schooling decisions and labor market experiences. Based on this information, individuals are assigned to either working in one of the three occupations, attending school, or simply staying at home.

Keane and Wolpin (1997) restrict attention to white males that turn 16 between 1977 and 1981 and exploit the information collected between 1979 and 1987. Thus, individuals in the sample are all between 16 and 26 years old. While the sample initially consists of 1,373 individuals at age 16, this number drops to 256 at the age of 26 due to sample attrition and missing data. Overall, the final sample consists of 12,359 person-period observations.

We briefly summarize the evolution of choices and wages over the sample period. Initially, roughly



(a) Choices

(b) Average wage

Notes: The wage is a full-time equivalent deflated by the gross national product deflator, with 1987 as the base year. We do not report the wage if less than ten observations are available.

Figure 6.: Data overview

86% of individuals enroll in school, but this share steadily declines with age. Nevertheless, about 39% obtain more than a high school degree and continue their schooling for more than twelve years. As individuals leave school, most of them initially pursue a blue-collar occupation. But the relative share of the white-collar occupation increases as individuals entering the labor market later have higher levels of schooling. At age 26, about 48% work in a blue-collar occupation and 34% in a white-collar occupation. The share of individuals in the military peaks around age 20 when it amounts to 8%. At its maximum around age 18, approximately 20% of individuals stay at home.

Overall, average wages start at about $10,000 at age 16 but increase considerably up to about $25,000 at age 26. While wages in the blue-collar occupation are initially highest with about $10,286, wages in the white-collar occupation and military start around $9,000. However, wages in the white-collar occupation increase steeper over time and overtake blue-collar wages around age 21. At the end of the observation period, wages in the white-collar occupation are about 50% higher than blue-collar wages with $32,756 as opposed to only $20,739. Military wages remain lowest throughout.

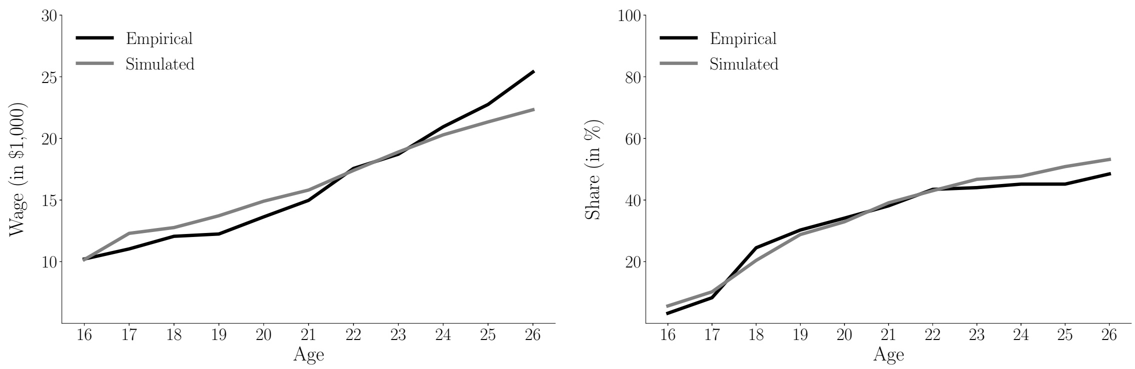
We consider observations for $i = 1, \ldots, N$ individuals in each time period $t = 1, \ldots, T_i$. For every observation (i, t) in the data, we observe the action a_{it}, some components $\bar{u}_{it}$ of the utility, and a subset $\bar{s}_{it}$ of the state s_{it}. Therefore, from an economist's point of view, we need to distinguish between two types of state variables $s_{it} = \{\bar{s}_{it}, \boldsymbol{e}, \boldsymbol{\epsilon_t}\}$. At time t, the economist and individual both observe $\bar{s}_{it}$ while $\{\boldsymbol{e}, \boldsymbol{\epsilon_t}\}$ is only observed by the individual.

We use simulated maximum likelihood (Fisher, 1922; Manski and Lerman, 1977) estimation and determine the 88 model parameters $\hat{\boldsymbol{\theta}}$ that maximize the likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathcal{D})$. As we only observe a subset $\bar{s}_t = \{\boldsymbol{k_t}, h_t, t, a_{t-1}\}$ of the state, we can determine the probability $p_{it}(a_{it}, \bar{u}_{it} \mid \bar{s}_{it}, \boldsymbol{\theta})$ of individual i at time t in $\bar{s}_{it}$ choosing a_{it} and receiving $\bar{u}_{it}$ given parametric assumptions about the distribution of $\boldsymbol{\epsilon_t}$. The objective function takes the following form:

$$\hat{\boldsymbol{\theta}} \equiv \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\max} \underbrace{\prod_{i=1}^{N} \prod_{t=1}^{T_i} p_{it}(a_{it}, \bar{u}_{it} \mid \bar{s}_{it}, \boldsymbol{\theta})}_{\mathcal{L}(\boldsymbol{\theta} \mid \mathcal{D})}.$$

Overall, our parameter estimates are in broad agreement with the results reported in the original paper and the related literature. For example, individuals discount future utilities by 6% per year. The returns to schooling differ by occupation. While wages in a white-collar occupation increase by about 6% with each year of schooling, they only increase by 2% in the blue collar occupation. Skills are transferable across occupations as both types of work experience increase wages in the blue and white-collar occupation.

Figure 7 shows the overall agreement between the empirical data and a dataset simulated using the estimated model parameters. We show average wages and the share of individuals choosing to work in a blue-collar occupation over time. The results are based on a simulated sample of 10,000 individuals. Additional model fit statistics are available in our Appendix.



(a) Average wage **(b)** Blue-collar

Figure 7.: Model fit

We follow the original authors and use the estimated model to conduct the ex-ante evaluation of a $2,000 tuition subsidy on educational attainment. We simulate a sample of 10,000 individuals using the point estimates and compare completed schooling to a sample of the same size but with a reduction of $\hat{\beta}_{tc_1}$ by $2,000. The subsidy increases average final schooling by 0.65 years. College graduation increases by 13 percentage points and high school graduation rates improve by 4 percentage points.

3.3. Confidence set bootstrap

The construction of confidence sets for counterfactuals in many structural models poses two distinct challenges. First, the computational burden of even a single estimation of the model is considerable. This makes the application of a standard bootstrap approach (Efron, 1979) infeasible. Second, the nonlinear mapping from the parameters of the model to the counterfactual predictions often has kinks or is truncated. For example, in our case the predicted impact of a tuition subsidy is bounded from below by zero. This violates the smoothness requirements of the delta method.

We use the Confidence Set (CS) bootstrap to construct the confidence set of the counterfactual. The CS bootstrap was originally proposed in Rao (1973) but only recently formalized by Woutersen and Ham (2019). Its application does not require repeated estimations of the model as it uses the asymptotic normal distribution of the estimator for $\hat{\boldsymbol{\theta}}$. Furthermore, its validity does not depend on the differentiability of the prediction function.[3]

Algorithm 1 provides a concise description of the steps involved, where $\chi^2_l(1-\alpha)$ is the quantile function for probability $1-\alpha$ of the chi-square distribution with l degrees of freedom.

Algorithm 1. . Confidence Set bootstrap

for $m = 1, \ldots, M$ **do**
  Draw $\hat{\boldsymbol{\theta}}_m \sim \mathcal{N}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$
  if $(\hat{\boldsymbol{\theta}}_m - \hat{\boldsymbol{\theta}})'\hat{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}_m - \hat{\boldsymbol{\theta}}) \leq \chi^2_l(1-\alpha)$ **then**
    Compute $\hat{y}_{g,m} = \mathcal{M}_g(\hat{\boldsymbol{\theta}}_m)$
    Add $\hat{y}_{g,m}$ to sample $Y = \{\hat{y}_{g,1}, \ldots, \hat{y}_{g,m-1}\}$
  end if
end for
Set $\Theta_{y_g}(\alpha) = [\min(Y), \max(Y)]$

In a nutshell, we draw a large sample of M parameters from the estimated asymptotic normal distribution of our estimator with mean $\hat{\boldsymbol{\theta}}$ and covariance matrix $\hat{\boldsymbol{\Sigma}}$, accept only those draws that are elements of the confidence set of the model parameters, compute the counterfactual for all remaining draws, and then calculate the confidence set for the counterfactual based on its lowest and highest

[3] See Reich and Judd (2020) for a critical assessment of confidence sets based on asymptotic arguments. They advocate the use of likelihood-ratio confidence intervals instead and cast their computation as a constraint optimization problem.

value.

The CS bootstrap poses a considerable computational challenge as in many applications, including ours, a single prediction of a counterfactual takes several minutes. At the same time, the number of parameter samples needs to be large to ensure that the minimum and maximum values for the counterfactual prediction are reliable. However, the algorithm is amenable to parallelization using modern high-performance computational resources as we can process each of the M parameter draws independently.

Our uncertainty sets then take the following form:

$$\begin{aligned}\mathcal{U}(\alpha) &\equiv \left\{\boldsymbol{\theta} \in \boldsymbol{\Theta} : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})'\hat{\Sigma}^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi^2_l(1-\alpha)\right\} \\ \mathcal{U}_{y_g}(\alpha) &\equiv \left\{M_g(\boldsymbol{\theta}) : (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})'\hat{\Sigma}^{-1}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi^2_l(1-\alpha), \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\}.\end{aligned}$$

4. Results

We now turn to the presentation of our results. Throughout, we focus on the impact of a $\$2,000$ tuition subsidy on completed schooling and use the 90% uncertainty set to measure the degree of uncertainty. All our results potentially depend on the size of the uncertainty set. In practice, policy-makers choose the uncertainty set's size in line with their underlying preferences - the more protection against unfavorable outcomes is desired, the larger the uncertainty set.[4]

All results are based on $30,000$ draws from the asymptotic normal distribution of our parameter estimates. We follow Keane and Wolpin (1997) and start by analyzing the prediction for a general subsidy. Then we turn to the situation where we use endowment types for policy targeting. We postulate a linear utility function for the policy-maker throughout.

4.1. General subsidy

Figure 8 explores the impact prediction for a general tuition subsidy. We show the point prediction, its sampling distribution, and the uncertainty set. At the point estimate, average schooling increases by 0.65 years. However, there is considerable uncertainty about the prediction as the uncertainty set ranges from 0.15 to 1.10 years.

[4]In a different setting, Blesch and Eisenhauer (2021) conduct an ex-ante performance evaluation of the statistical decision functions over the whole parameter space (Wald, 1950; Manski, 2021).

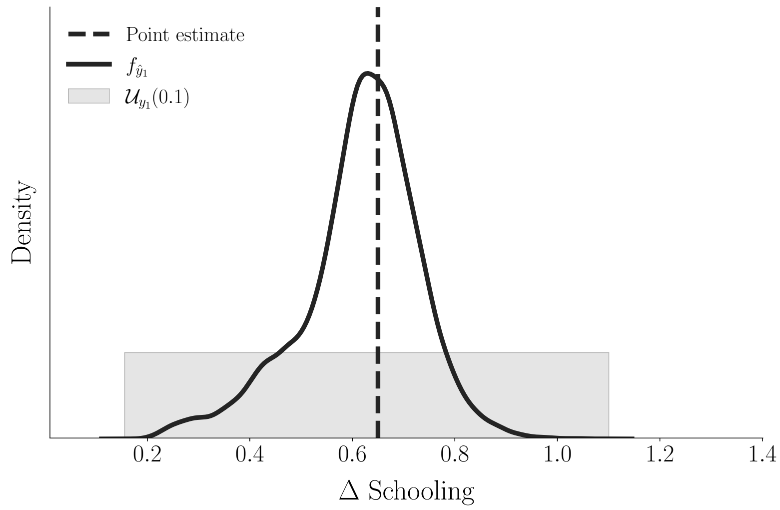


Figure 8.: General subsidy

In Figure 9, we trace out the effect of the discount rate δ on the subsidy's impact over the uncertainty set while keeping all other parameters at their point estimate. Initially, as δ increases, so does the policy's impact as individuals value the long-term benefits from increasing their level of schooling more and more. However, for high levels of the discount factor, the policy's impact starts to decrease as most individuals already complete a high school or college degree even without the subsidy.

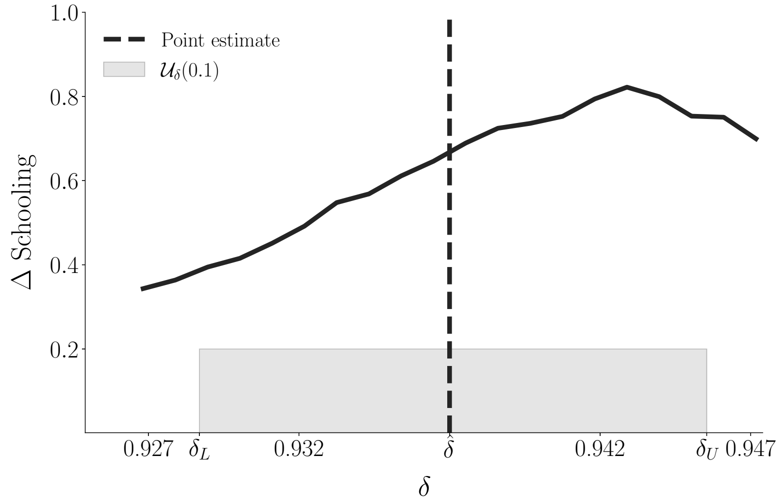


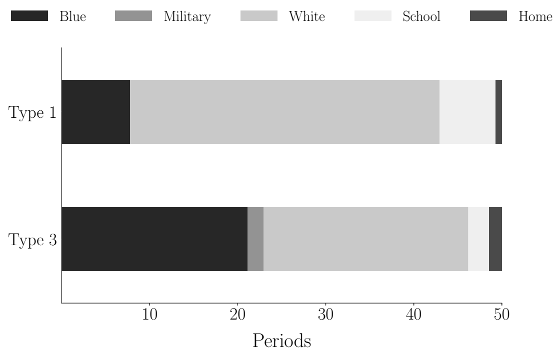
Figure 9.: Time preference

4.2. Targeted subsidy

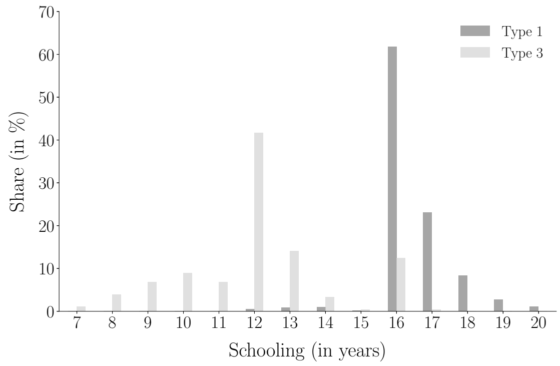
So far, we restricted the analysis to a general subsidy available to the whole population and the average predicted impact. We now move to the setting where a policy-maker can target individuals by their type of initial endowment. The importance of early endowment heterogeneity in shaping economic outcomes over the life-cycle is the most important finding from Keane and Wolpin (1997). It served as motivation for a host of subsequent research on the determinants of skill heterogeneity among adolescents (Todd and Wolpin, 2007; Erosa et al., 2010; Caucutt and Lochner, 2020).

To ease the exposition, we initially focus our discussion of results on Type 1 and Type 3 individuals. We then later rank policies targeting either of the four types based on the different decision-theoretic criteria. Additional results are available in our Appendix.

Figure 10 confirms that life-cycle choices differ considerably by initial endowment type. On the left, we show the number of periods the two types spend on average in each of the five alternatives. Type 1 spends more than six years in education even after entering the model. Type 3 only adds another two years. This difference translates into very different labor market experiences. While Type 1 works for about 35 years in a white-collar occupation, Type 3 switches more frequently and spends an equal number of 22 years in a white and a blue-collar occupation. Both types only spend a short time at home.



(a) Life-cycle choices

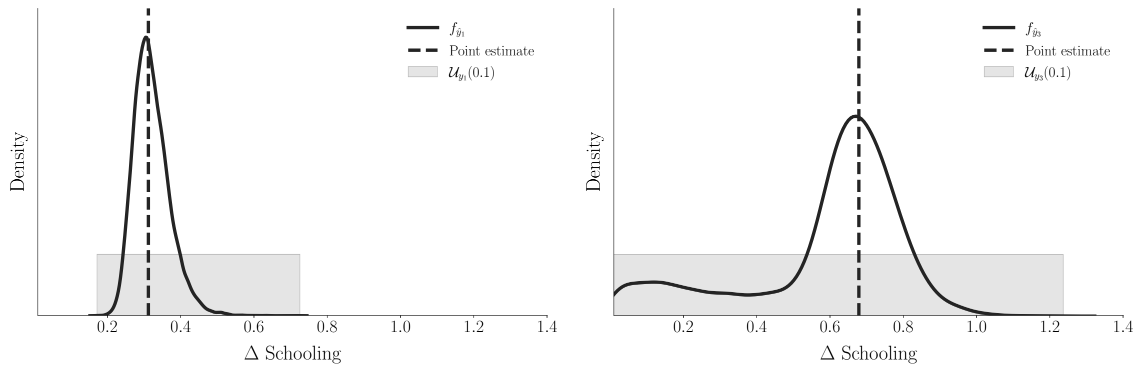


(b) Completed schooling

Figure 10.: Type heterogeneity

On the right, we show the distribution of final schooling for the two types. Average schooling is considerably higher for Type 1 with more than 16 years, compared to about 12 years for Type 3. Nearly all Type 1 individuals enroll in college, and most do end up with a degree.

Figure 11 provides a visualization of our core results for a targeted subsidy. At the point estimates, the predicted impact is considerably lower for Type 1 than Type 3. However, the prediction uncertainty is much larger for Type 3 compared to Type 1. The uncertainty set for Type 3 ranges all the way from 0 to 1.2 years, while the prediction for Type 1 is between 0.18 and 0.75.



(a) Type 1

(b) Type 3

Figure 11.: Targeted subsidy

This heterogeneity in impact and prediction uncertainty follows directly from the underlying economics of the model. Type 1 is already much more likely to have a college degree before the subsidy, and thus, the predicted impact is smaller. At the same time, Type 1 individuals affected by the subsidy are already pursuing some college and thus directly benefit from the subsidy. Type 3 is at the lower end of the schooling distribution. A tuition subsidy can considerably increase their level of schooling, but whether the subsidy succeeds in doing so is uncertain.

We now consider the policy option to target Type 2 and Type 4 as well. Their point predictions are actually highest with an additional 0.81 years on average for Type 2 and 0.75 years for Type 4. However, both predictions are fraught with uncertainty. For Type 2 the uncertainty set ranges from 0.17 to 1.3, while for Type 4 it starts at zero and spans all the way to 1.18.

Figure 12 shows the policy alternative's ranking by the decision-theoretic criteria we discussed in Section 2.2. Ranking alternatives using as-if optimization is straightforward. A policy targeting Type 2 is the most preferred alternative, while a focus on Type 1 is the least attractive. However, once we take the uncertainty in the predictions into account, a more nuanced picture emerges. Moving from as-if optimization to a subjective Bayes criterion using a uniform distribution over the uncertainty set does not change the ordering. However, once a decision-maker is concerned with performance across the whole range of values in the uncertainty set – we move to the minimax regret or maximin criterion – a policy targeting Type 1 becomes more and more attractive despite its low point prediction because its worst-case utility is highest.

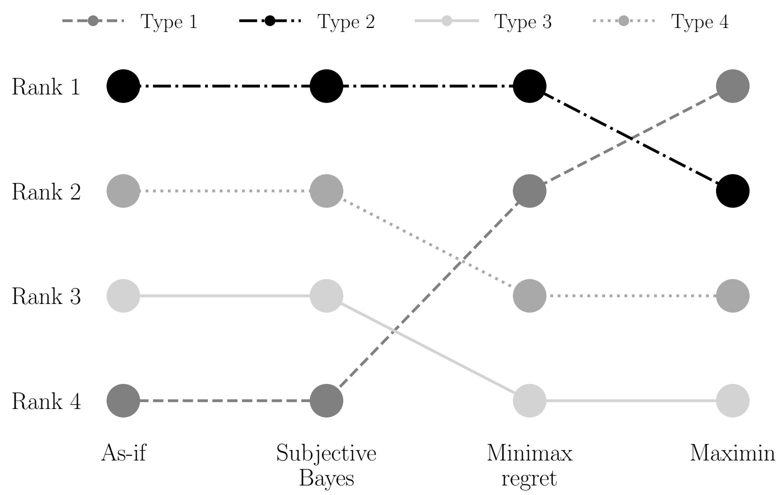


Figure 12.: Policy ranking

In general, framing policy advice as a decision problem under uncertainty shows that there are many different ways of making reasonable decisions. Different criteria result in different policy rankings. Not only that, the ranking of policies for a given criteria potentially depends on the choice of α. We think of α as part of a policy-maker's decision problem: The more weight the policy-maker places on worst-case events, the smaller the appropriate value for α. A policy-maker should decide on their preferred decision rule and perform a sensitivity analysis around the selected α values. This approach allows learning about yet another layer of uncertainty concerning the preferred policy choice.

5. Conclusion

We develop a generic approach that deals with parametric uncertainty when using models to inform policy-making. We propose a decision-theoretic analysis of even computationally demanding structural models based on uncertainty sets. We construct the uncertainty sets from empirical estimates and ensure their computational tractability using the Confidence Set bootstrap. We revisit the seminal work by Keane and Wolpin (1997) to document the empirical relevance of prediction uncertainty and showcase our analysis. Focusing on their ex-ante evaluation of a tuition subsidy, we report considerable uncertainty in its impact on completed schooling. We show how a policy-maker's preferred policy depends on the choice of alternative formal rules for decision-making under uncertainty.

In ongoing work, we pursue three avenues for further improvements. First, we link our work with the literature on inference under (local) model misspecification to refine the construction of our uncertainty sets. For example, Armstrong and Kolesár (2021) and Bonhomme and Weidner (2020) propose different methods for taking misspecification into account when constructing confidence sets. Second, we incorporate ideas from the literature on global sensitivity analysis (Razavi et al., 2021) to identify the parameters most responsible for the uncertainty in predictions. The attribution of importance based on Shapely values, familiar to economists from game theory, appears promising (Shapley, 1953; Owen, 2014). Third, we address our analysis's computational burden using surrogate modeling (Forrester et al., 2008). A surrogate model emulates the full model's behavior at a negligible cost per run and allows us to determine prediction uncertainty using a nonparametric bootstrap procedure.

References

Adda, J., Dustmann, C., and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2):293–337.

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67.

Altonji, J. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 11(1):48–83.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2020). On the informativeness of descriptive statistics for structural estimates. *Econometrica*, 88(6):2231–2258.

Armstrong, T. B. and Kolesár, M. (2021). Sensitivity analysis using approximate moment condition models. *Quantitative Economics*, 12(1):77–108.

Ben-Tal, A., den Hertog, D., De Waegenaere, A., Melenberg, B., and Rennen, G. (2013). Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357.

Berger, L., Berger, N., Bosetti, V., Gilboa, I., Hansen, L. P., Jarvis, C., Marinacci, M., and Smith, R. D. (2020). Rational policymaking during a pandemic. *Proceedings of the National Academy of Sciences*, 118(4).

Bertsimas, D., Gupta, V., and Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming*, 167(2):235–292.

Blesch, M. and Eisenhauer, P. (2021). Robust investments under risk and ambiguity. *under revision*.

Blundell, R. (2017). What have we learned from structural models? *American Economic Review*, 107(5):287–92.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Blundell, R. and Shephard, A. (2012). Employment, hours of work and the optimal taxation of low-income families. *The Review of Economic Studies*, 79(2):481–510.

Bonhomme, S. and Weidner, M. (2020). Minimizing sensitivity to model misspecification. *arXiv Working Paper*.

Bugni, F. and Ura, T. (2019). Inference in dynamic discrete choice problems under local misspecification. *Quantitative Economics*, 10(1):67–103.

Bureau of Labor Statistics (2019). *National Longitudinal Survey of Youth 1979 cohort, 1979-2016 (rounds 1-27)*. Center for Human Resource Research, Ohio State University, Columbus, OH.

Cai, Y. and Lontzek, T. S. (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy*, 127(6):2684–2734.

Caucutt, E. M. and Lochner, L. (2020). Early and late human capital investments, borrowing constraints, and the family. *Journal of Political Economy*, 128(3):1065–1147.

Christensen, T. and Connault, B. (2019). Counterfactual sensitivity and robustness. *arXiv Working Paper*.

Cunha, F., Heckman, J. J., and Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Eckstein, Z., Keane, M., and Lifshitz, O. (2019). Career and family decisions: Cohorts born 1935-1975. *Econometrica*, 87(1):217–253.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26.

Eisenhauer, P., Heckman, J. J., and Mosso, S. (2015). Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review*, 56(2):331–357.

Erosa, A., Koreshkova, T., and Restuccia, D. (2010). How important is human capital? A quantitative theory assessment of world income inequality. *The Review of Economic Studies*, 77(4):1421–1449.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222(594-604):309–368.

Forrester, D. A. I. J., Sobester, D. A., and Keane, P. A. J. (2008). *Engineering design via surrogate modelling: A practical guide*. John Wiley & Sons, Chichester, England.

Gabler, J. and Raabe, T. (2020). respy - a framework for the simulation and estimation of Eckstein-Keane-Wolpin models.

Gayle, G.-L. and Shephard, A. (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica*, 87(1):291–326.

Gilboa, I. (2009). *Theory of decision under uncertainty*. Cambridge University Press, New York City, NY.

Gilboa, I. (2020). What were you thinking? Revealed preference theory as coherence test. *Working Paper*.

Gilboa, I., Rouziou, M., and Sibony, O. (2018). Decision theory made relevant: Between the software and the shrink. *Research in Economics*, 72(2):240–250.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Harenberg, D., Marelli, S., Sudret, B., and Winschel, V. (2019). Uncertainty quantification and global sensitivity analysis for economic models. *Quantitative Economics*, 10(1):1–41.

Hood, W. and Koopmans, T. (1953). *Studies in econometric method*. John Wiley & Sons, New York City, NY.

Hungerford, T. and Solon, G. (1987). Sheepskin effects in the returns to education. *Review of Economics and Statistics*, 69(1):175–177.

Jaeger, D. A. and Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics*, 78(4):733–740.

Jørgensen, T. H. (2021). Sensitivity to calibrated paramters. *Review of Economics and Statistics*, forthcoming.

Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2):375–406.

Keane, M. P., Todd, P. E., and Wolpin, K. I. (2011). The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, pages 331–461. Elsevier Science, Amsterdam, Netherlands.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.

Low, H. and Meghir, C. (2017). The use of structural models in econometrics. *Journal of Economic Perspectives*, 31(2):33–58.

Machina, M. J. and Viscusi, K. (2014). *Handbook of the economics of risk and uncertainty.* North-Holland Publishing Company, Amsterdam, Netherlands.

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica*, 72(4):1221–1246.

Manski, C. F. (2013). *Public policy in an uncertain world: Analysis and decisions.* Harvard University Press, Cambridge, MA.

Manski, C. F. (2019). Communicating uncertainty in policy analysis. *Proceedings of the National Academy of Sciences*, 116(16):7634–7641.

Manski, C. F. (2021). Econometrics for decision making: Building foundations sketched by Haavelmo and Wald. arxiv Working Paper, National Bureau of Economic Research.

Manski, C. F. and Lerman, S. R. (1977). The estimation of choice probabilities from choice based samples. *Econometrica*, 45(8):1977–1988.

Marinacci, M. (2015). Model uncertainty. *Journal of the European Economic Association*, 13(6):1022–1100.

Mincer, J. (1974). *Schooling, experience and earnings.* National Bureau of Economic Research, New York City, NY.

Mukhin, Y. (2018). Sensitivity of regular estimators. *arXiv Working Paper.*

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3):315–335.

National Research Council (2012). *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification.* The National Academies Press, Washington, DC.

Niehans (1948). Zur Preisbildung bei ungewissen Erwartungen. *Swiss Journal of Economics and Statistics*, 84(5):433–456.

Owen, A. B. (2014). Sobol' indices and shapley value. *Journal on Uncertainty Quantification*, 2(1):245–251.

Puterman, M. L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, New York City, NY.

Rao, C. R. (1973). *Linear statistical inference and its applications*. John Wiley & Sons, New York City, NY.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Piano, S. L., Iwanaga, T., Becker, W., et al. (2021). The future of sensitivity analysis: An essential discipline for systems modeling and policy support. *Environmental Modeling & Software*, 137:104954.

Reich, G. and Judd, K. L. (2020). Efficient likelihood ratio confidence intervals using constrained optimization. *SSRN Working Paper*.

Rust, J. (1994). Structural estimation of Markov decision processes. In Engle, R. and McFadden, D., editors, *Handbook of Econometrics*, pages 3081–3143. North-Holland Publishing Company, Amsterdam, Netherlands.

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Espeland, W. N., Kay, J., Piano, S. L., Mayo, D., et al. (2020). Five ways to ensure that models serve society: A manifesto. *Nature*, 582:482–484.

Samuelson, P. A. (1937). A note on measurement of utility. *Review of Economic Studies*, 4(2):155–161.

SAPEA (2019). Making sense of science for policy under conditions of complexity and uncertainty.

Savage, L. J. (1954). *The foundations of statistics*. John Wiley & Sons, New York City, NY.

Shapley, L. S. (1953). *A value for n-person games*. Princeton University Press, Princeton, NJ.

Todd, P. E. and Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5):1384–1417.

Todd, P. E. and Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital*, 1(1):91–136.

Wald, A. (1950). *Statistical decision functions.* John Wiley & Sons, New York City, NY.

White, D. J. (1993). *Markov decision processes.* John Wiley & Sons, New York City, NY.

Wolpin, K. I. (2013). *The limits to inference without theory.* MIT University Press, Cambridge, MA.

Woutersen, T. and Ham, J. (2019). Confidence sets for continuous and discontinuous functions of parameters. *SSRN Working Paper*.

A. Appendix

Our Appendix contains details on our computational implementation, information about the estimation dataset, and additional results.

A.1. Computation

We use the same computational implementation as in Keane and Wolpin (1997). We outline the immediate utility functions for each of the five alternatives. We first focus on their common overall structure and then present their parameterization. Throughout we provide the economic motivation for their specification.

We follow individuals over their working life from young adulthood at age 16 to retirement at age 65. The decision period $t = 16, \ldots, 65$ is a school year, and individuals decide $a \in \mathcal{A}$ whether to work in a blue-collar or white-collar occupation ($a = 1, 2$), to serve in the military ($a = 3$), to attend school ($a = 4$), or to stay at home ($a = 5$).

Individuals are initially heterogeneous. They differ with respect to their initial level of completed schooling h_{16} and have one of four different $\mathcal{J} = \{1, \ldots, 4\}$ alternative-specific skill endowments $\boldsymbol{e} = (e_{j,a})_{\mathcal{J}\times\mathcal{A}}$.

The immediate utility $u_a(\cdot)$ of each alternative consists of a non-pecuniary utility $\zeta_a(\cdot)$ and, at least for the working alternatives, an additional wage component $w_a(\cdot)$. Both depend on the level of human capital as measured by their occupation-specific work experience $\boldsymbol{k}_t = (k_{a,t})_{a\in\{1,2,3\}}$, years of completed schooling h_t, and alternative-specific skill endowment $\boldsymbol{e}$. The immediate utility functions are influenced by last-period choices a_{t-1} and alternative-specific productivity shocks $\boldsymbol{\epsilon}_t = (\epsilon_{a,t})_{a\in\mathcal{A}}$ as well. Their general form is given by:

$$u_a(\cdot) = \begin{cases} \zeta_a(\boldsymbol{k}_t, h_t, t, a_{t-1}) + w_a(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if}\, a \in \{1, 2, 3\} \\ \zeta_a(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) & \text{if}\, a \in \{4, 5\}. \end{cases}$$

Work experience $\boldsymbol{k}_t$ and years of completed schooling h_t evolve deterministically:

$$\begin{aligned} k_{a,t+1} &= k_{a,t} + \mathbf{I}[a_t = a] \qquad \text{if}\, a \in \{1, 2, 3\} \\ h_{t+1} &= h_t \;\;\, + \mathbf{I}[a_t = 4]. \end{aligned}$$

The productivity shocks are uncorrelated across time and follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$. Given the structure of the utility functions and the distribution of the shocks, the state at time t is $s_t = \{\boldsymbol{k}_t, h_t, t, a_{t-1}, \boldsymbol{e}, \boldsymbol{\epsilon}_t\}$.

Empirical and theoretical research from specialized disciplines within economics informs the exact specification of $u_a(\cdot)$. We now discuss each of its components in detail.

Non-pecuniary utility

We present the parameterization of the non-pecuniary utility for all five alternatives.

Blue-collar Equation (3) shows the parameterization of the non-pecuniary utility from working in a blue-collar occupation:

$$\begin{aligned}\zeta_1(\boldsymbol{k}_t, h_t, a_{t-1}) = \alpha_1 + c_{1,1} \cdot \mathbf{I}[a_{t-1} \neq 1] + c_{1,2} \cdot \mathbf{I}[k_{1,t} = 0] \\ + \vartheta_1 \cdot \mathbf{I}[h_t \geq 12] + \vartheta_2 \cdot \mathbf{I}[h_t \geq 16] + \vartheta_3 \cdot \mathbf{I}[k_{3,t} = 1].\end{aligned} \tag{3}$$

A constant α_1 captures the net monetary-equivalent of on the job amenities. The non-pecuniary utility includes mobility and search costs $c_{1,1}$, which are higher for individuals who never worked in a blue-collar occupation before $c_{1,2}$. The non-pecuniary utilities capture returns from a high school ϑ_1 and a college ϑ_2 degree. Additionally, there is a detrimental effect of leaving the military early after one year ϑ_3.

White-collar The non-pecuniary utility from working in a white-collar occupation is specified analogously. Equation (4) shows its parameterization:

$$\begin{aligned}\zeta_2(\boldsymbol{k}_t, h_t, a_{t-1}) = \alpha_2 + c_{2,1} \cdot \mathbf{I}[a_{t-1} \neq 2] + c_{2,2} \cdot \mathbf{I}[k_{2,t} = 0] \\ + \vartheta_1 \cdot \mathbf{I}[h_t \geq 12] + \vartheta_2 \cdot \mathbf{I}[h_t \geq 16] + \vartheta_3 \cdot \mathbf{I}[k_{3,t} = 1].\end{aligned} \tag{4}$$

Military Equation (5) shows the parameterization of the non-pecuniary utility from working in the military:

$$\zeta_3(k_{3.t}, h_t) = c_{3,2} \cdot \mathbf{I}[k_{3,t} = 0] + \vartheta_1 \cdot \mathbf{I}[h_t \geq 12] + \vartheta_2 \cdot \mathbf{I}[h_t \geq 16]. \tag{5}$$

Search costs $c_{3,1} = 0$ are absent but there is a mobility cost if an individual has never served in the military before $c_{3,2}$. Individuals still experience a non-pecuniary utility from finishing high-school ϑ_1 and college ϑ_2.

School Equation (6) shows the parameterization of the non-pecuniary utility from schooling:

$$\begin{aligned}\zeta_4(k_{3,t}, h_t, t, a_{t-1}, e_{j,4}, \epsilon_{4,t}) = e_{j,4} &+ \beta_{tc_1} \cdot \mathbf{I}[h_t \geq 12] + \beta_{tc_2} \cdot \mathbf{I}[h_t \geq 16] \\ &+ \beta_{rc_1} \cdot \mathbf{I}[a_{t-1} \neq 4, h_t < 12] \\ &+ \beta_{rc_2} \cdot \mathbf{I}[a_{t-1} \neq 4, h_t \geq 12] + \gamma_{4,4} \cdot t \\ &+ \gamma_{4,5} \cdot \mathbf{I}[t < 18] + \vartheta_1 \cdot \mathbf{I}[h_t \geq 12] \\ &+ \vartheta_2 \cdot \mathbf{I}[h_t \geq 16] + \vartheta_3 \cdot \mathbf{I}[k_{3,t} = 1] + \epsilon_{4,t}.\end{aligned} \tag{6}$$

There is a direct cost of attending school such as tuition for continuing education after high school β_{tc_1} and college β_{tc_2}. The decision to leave school is reversible, but entails adjustment costs that differ by schooling category $(\beta_{rc_1}, \beta_{rc_2})$. Schooling is defined as time spent in school and not by formal credentials acquired. Once individuals reach a certain amount of schooling, they acquire a degree. There is no uncertainty about grade completion (Altonji, 1993) and no part-time enrollment. Individuals value the completion of high-school and graduate school $(\vartheta_1, \vartheta_2)$.

Home Equation (7) shows the parameterization of the non-pecuniary utility from staying at home:

$$\begin{aligned}\zeta_5(k_{3,t}, h_t, t, e_{j,5}, \epsilon_{5,1}) = e_{j,5} &+ \gamma_{5,4} \cdot \mathbf{I}[18 \leq t \leq 20] + \gamma_{5,5} \cdot \mathbf{I}[t \geq 21] \\ &+ \vartheta_1 \cdot \mathbf{I}[h_t \geq 12] + \vartheta_2 \cdot \mathbf{I}[h_t \geq 16] \\ &+ \vartheta_3 \cdot \mathbf{I}[k_{3,t} = 1] + \epsilon_{5,t}.\end{aligned} \tag{7}$$

Staying at home as a young adult $\gamma_{5,4}$ is less stigmatic as doing so while already being an adult $\gamma_{5,5}$. Additionally, possessing a degree $(\vartheta_1, \vartheta_2)$ or leaving the military prematurely ϑ_3 influences the immediate utility.

Wage component

The wage component $w_a(\cdot)$ for the working alternatives is given by the product of the market-equilibrium rental price r_a and an occupation-specific skill level $x_a(\cdot)$. The latter is determined by the overall level of human capital:

$$w_a(\cdot) = r_a \, x_a(\cdot).$$

This specification leads to a standard logarithmic wage equation in which the constant term is the skill rental price $\ln(r_a)$ and wages follow a log-normal distribution.

The occupation-specific skill level $x_a(\cdot)$ is determined by a skill production function, which includes a deterministic component $\Gamma_a(\cdot)$ and a multiplicative stochastic productivity shock $\epsilon_{a,t}$:

$$x_a(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,a}, \epsilon_{a,t}) = \exp\big(\Gamma_a(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,a}) \cdot \epsilon_{a,t}\big).$$

Blue-collar Equation (8) shows the parameterization of the deterministic component of the skill production function:

$$\begin{aligned}\Gamma_1(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,1}) = e_{j,1} &+ \beta_{1,1} \cdot h_t + \beta_{1,2} \cdot \mathbf{I}[h_t \geq 12] \\ &+ \beta_{1,3} \cdot \mathbf{I}[h_t \geq 16] + \gamma_{1,1} \cdot k_{1,t} + \gamma_{1,2} \cdot (k_{1,t})^2 \\ &+ \gamma_{1,3} \cdot \mathbf{I}[k_{1,t} > 0] + \gamma_{1,4} \cdot t + \gamma_{1,5} \cdot \mathbf{I}[t < 18] \\ &+ \gamma_{1,6} \cdot \mathbf{I}[a_{t-1} = 1] + \gamma_{1,7} \cdot k_{2,t} + \gamma_{1,8} \cdot k_{3,t}.\end{aligned} \tag{8}$$

There are several notable features. The first part of the skill production function is motivated by Mincer (1974) and hence linear in years of completed schooling $\beta_{1,1}$, quadratic in experience $(\gamma_{1,1}, \gamma_{1,2})$, and separable between the two of them. There are so-called sheep-skin effects (Hungerford and Solon, 1987; Jaeger and Page, 1996) associated with completing a high school $\beta_{1,2}$ and graduate $\beta_{1,3}$ education that capture the impact of completing a degree beyond just the associated years of schooling. Also, there is a first-year blue-collar experience effect $\gamma_{1,3}$ while skills depreciate when not employed in a blue-collar occupation in the preceding period $\gamma_{1,6}$. Other work experience $(\gamma_{1,7}, \gamma_{1,8})$ is transferable.

White-collar The wage component from working in a white-collar occupation is specified analogously. Equation (9) shows the parameterization of the deterministic component of the skill production function:

$$\begin{aligned}\Gamma_2(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,2}) = e_{j,2} &+ \beta_{2,1} \cdot h_t + \beta_{2,2} \cdot \mathbf{I}[h_t \geq 12] \\ &+ \beta_{2,3} \cdot \mathbf{I}[h_t \geq 16] + \gamma_{2,1} \cdot k_{2,t} + \gamma_{2,2} \cdot (k_{2,t})^2 \\ &+ \gamma_{2,3} \cdot \mathbf{I}[k_{2,t} > 0] + \gamma_{2,4} \cdot t + \gamma_{2,5} \cdot \mathbf{I}[t < 18] \\ &+ \gamma_{2,6} \cdot \mathbf{I}[a_{t-1} = 2] + \gamma_{2,7} \cdot k_{1,t} + \gamma_{2,8} \cdot k_{3,t}.\end{aligned} \tag{9}$$

Military Equation (10) shows the parameterization of the deterministic component of the skill production function:

$$\begin{aligned}\Gamma_3(k_{3,t}, h_t, t, e_{j,3}) = e_{j,3} &+ \beta_{3,1} \cdot h_t \\ &+ \gamma_{3,1} \cdot k_{3,t} + \gamma_{3,2} \cdot (k_{3,t})^2 + \gamma_{3,3} \cdot \mathbf{I}[k_{3,t} > 0] \\ &+ \gamma_{3,4} \cdot t + \gamma_{3,5} \cdot \mathbf{I}[t < 18].\end{aligned} \tag{10}$$

Contrary to the civilian sector there are no sheep-skin effects from graduation ($\beta_{3,2} = \beta_{3,3} = 0$). The previous occupational choice has no influence ($\gamma_{3,6} = 0$) and any experience other than military is non-transferable ($\gamma_{3,7} = \gamma_{3,8} = 0$).

Remark 1 *Our parameterization for the immediate utility of serving in the military differs from Keane and Wolpin (1997) as we remain unsure about their exact specification. The authors state in Footnote 31 (p. 498) that the constant for the non-pecuniary utility $\alpha_{3,t}$ depends on age. However, we are unable to determine the precise nature of the relationship. Equation (C3) (p. 521) also indicates no productivity shock $\epsilon_{a,t}$ in the wage component. Table 7 (p. 500) reports such estimates.*

A.2. Data

We use the same data as in Keane and Wolpin (1997). They construct their sample based on the National Longitudinal Survey of Youth 1979 (NLSY79) (Bureau of Labor Statistics, 2019). The NLSY79 is a nationally representative sample of young men and women living in the United States in 1979 and born between 1957 and 1964. Individuals were followed from 1979 onwards and repeatedly interviewed about their educational decisions and labor market experiences. Based on this information, individuals are assigned to either working in one of the three occupations, attending school, or simply staying at home. The decision period is the school year. The sample is restricted to white males that turn 16 between 1977 and 1981 and uses information collected between 1979 and 1987. Thus individuals in the sample are all between 16 and 26 years old.

Figure 13 shows the sample size by age. While the sample initially consists of 1,373 individuals at age 16, this number drops to 256 at the age of 26 due to sample attrition, missing data, and the short observation period. Overall, the final sample consists of 12,359 person-period observations.

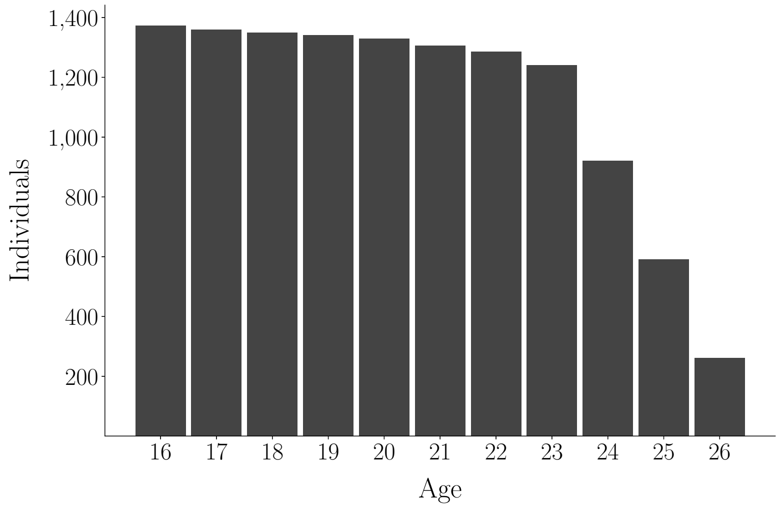


Figure 13.: Sample size

Figure 14 shows the distribution of initial schooling among individuals when entering the model. The large majority of individuals enter the model with ten years of schooling, while about a quarter of individuals has less than ten years of schooling. About 7.5% of individuals already attended school for 11 years.

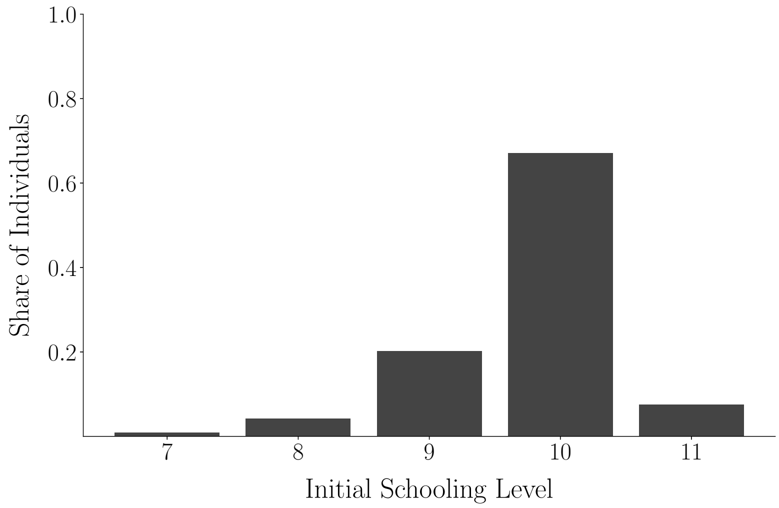


Figure 14.: Initial schooling

Figure 15 shows heterogeneity of choices by the level of initial schooling. Individuals who enter the model with only seven years of schooling spend another 0.65 years in school after age 16. Consequently, they spend around four years at home and, if they are working, then do so in a blue-collar occupation. When starting with ten years of schooling, then individuals add roughly another three years while in the model. This increase is about half a year more than individuals that start with eleven years.

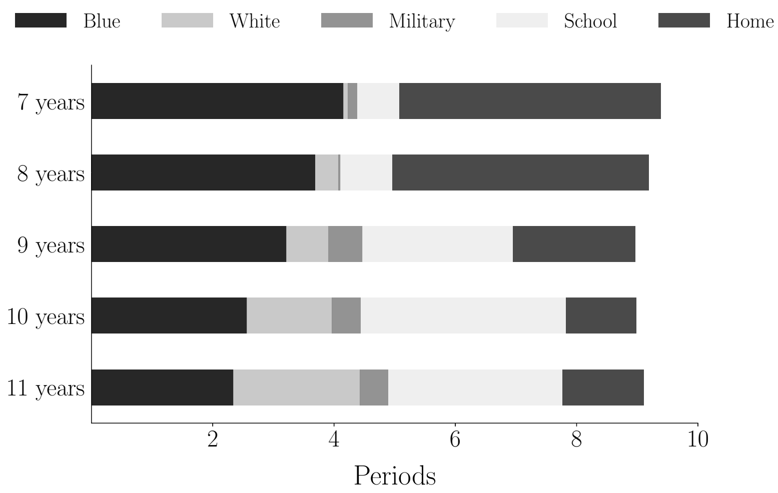


Figure 15.: Average choices by initial schooling

Figure 16 documents strong persistence in choices over time. For example, among those that work in a white-collar occupation in t 67% work in the same occupation in $t+1$ and 20% switch to a blue-collar occupation.

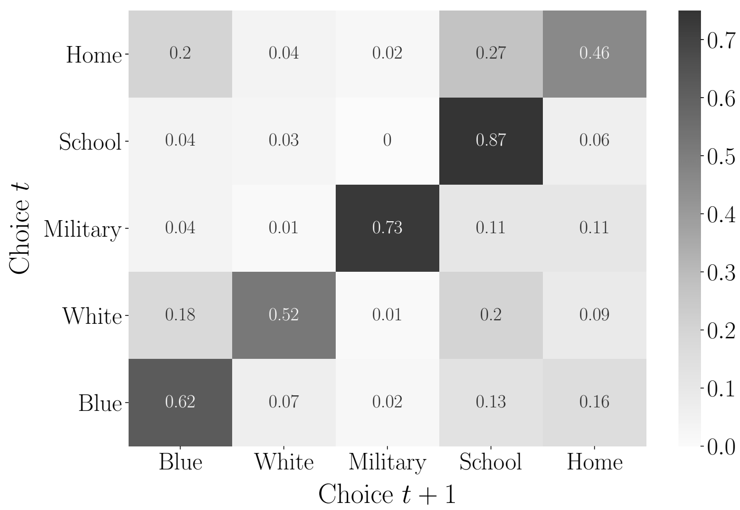
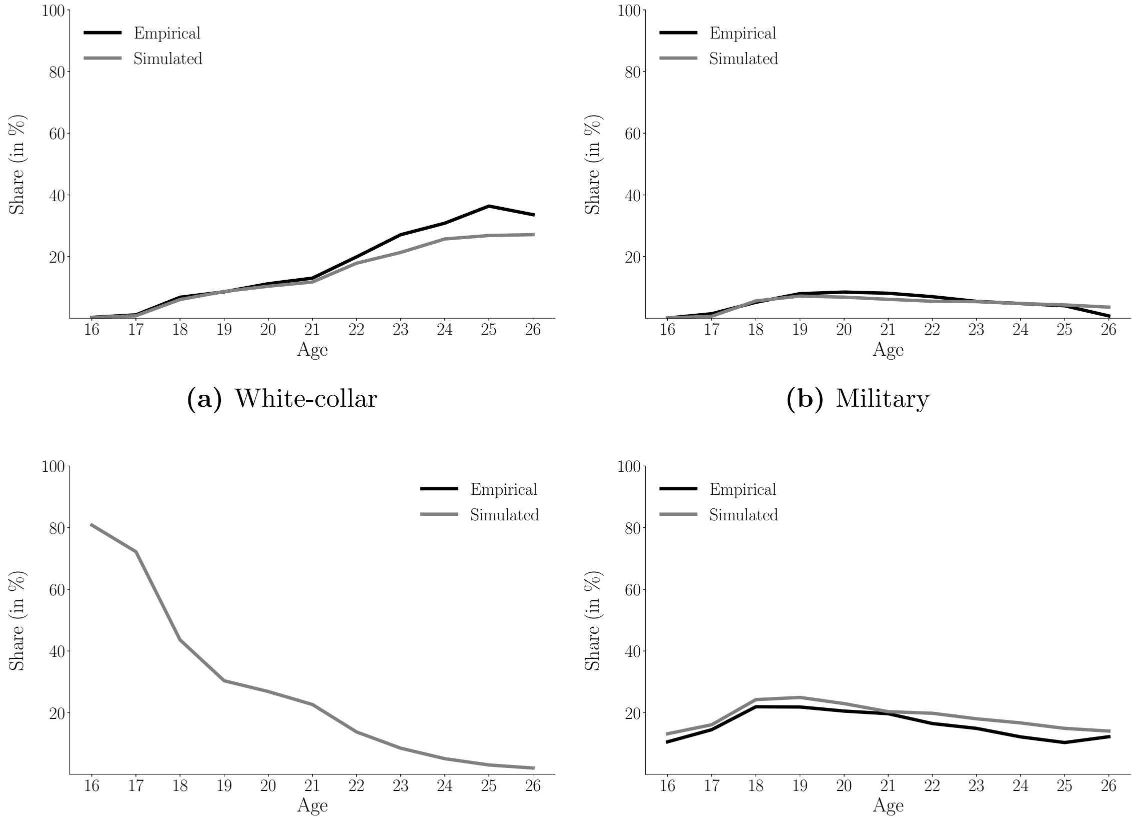


Figure 16.: Transition matrix

A.3. Results

Figure 17 shows further comparisons between the simulated and empirical data. All results are based on $10,000$ individuals using the estimated model.



(a) White-collar

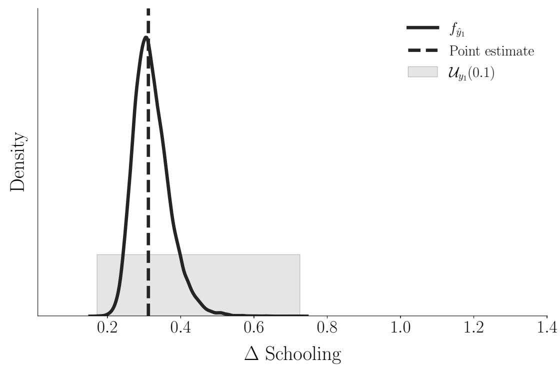
(b) Military

(c) School

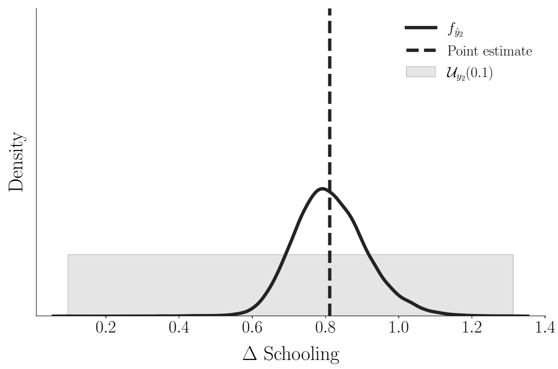
(d) Home

Figure 17.: Model fit

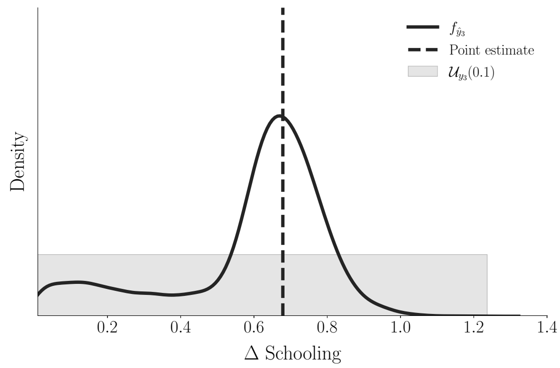
Figure 18 provides the point prediction, its sampling distribution, and the estimated confidence set for the impact of the tuition subsidy for all types. All results are based on $30,000$ draws from the asymptotic normal distribution of our parameter estimates.



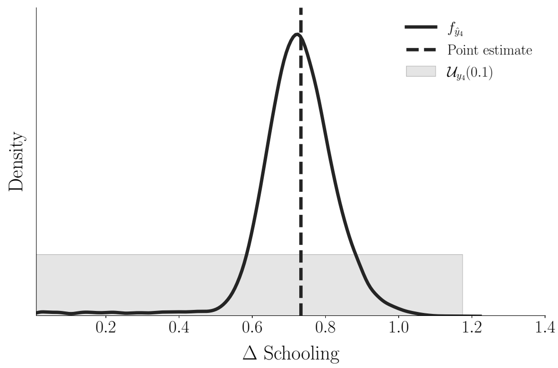
(a) Type 1



(b) Type 2



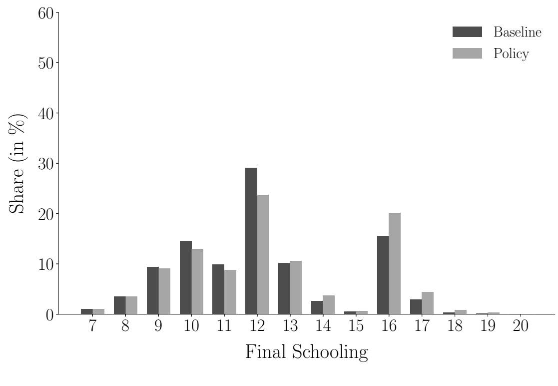
(c) Type 3



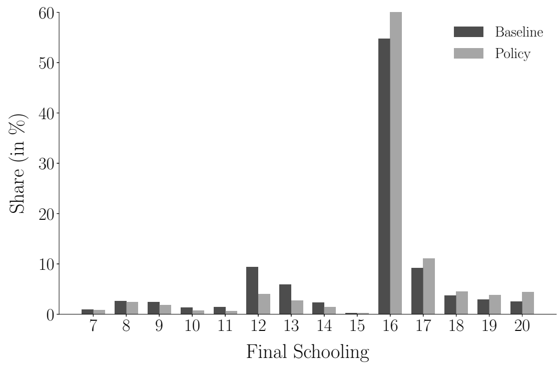
(d) Type 4

Figure 18.: Targeted subsidy for all types

Figure 19 shows the impact of the tuition subsidy at the upper δ_H and lower δ_L bound of the estimated confidence set for δ. The results are based on simulated samples of $10,000$ individuals for both scenarios.



(a) Low δ



(b) High δ

Figure 19.: Policy impact and time preference