# Eckstein–Keane–Wolpin models
## An invitation for transdisciplinary collaboration*

Open Source Economics

This version: November 10, 2020


**Abstract**

We present background material on a class of structural microeconometric models to facilitate transdisciplinary collaboration in their future development. We describe the economic framework, mathematical formulation, and calibration procedures for the so-called Eckstein–Keane–Wolpin (EKW) models. We provide an exemplifying analysis of the seminal model outlined in Keane and Wolpin (1997) and present our group's ensemble of research codes that allow for its specification, simulation, and calibration. We summarize our efforts drawing on research outside economics to address the computational challenges in applying EKW models and improve the reliability and interpretability of their results.



*Corresponding author: Philipp Eisenhauer, p.eisenhauer@uni-bonn.de.


We thank Janoś Gabler, Annica Gehlen, Boryana Ilieva, Fedor Ishakov, Lena Janys, Moritz Mendel, Tim Mensinger, Tobias Raabe, Klara Röhrl, John Rust, and Rafael Suchy for their valuable input. Luis Wardenbach provided outstanding research assistance.

## Contents

# 1 Introduction

Economists use structural microeconometric models to study individual decision-making. These models specify the objective of individuals, their economic environment, and the institutional and informational constraints under which they operate. Calibration of the model to observed data on individual decisions and experiences allows quantifying the importance of competing economic mechanisms in determining economic outcomes and forecasting the effects of policy proposals (Wolpin, 2013).

We restrict our exposition to the class of Eckstein–Keane–Wolpin (EKW) models (Adda et al., 2017; Blundell et al., 2016; Keane & Wolpin, 1997). Labor economists use them to study human capital investment decisions. Human capital comprises the knowledge, skills, competencies, and attributes embodied in individuals facilitating the creation of personal, social, and economic well-being (Becker, 1964). Differences in human capital attainment lead to inequality in various life outcomes such as labor market success and health across and within countries (OECD, 2001).

In Bhuller et al. (2018), for example, we apply an EKW model to analyze the mechanisms determining schooling decisions in Norway. We calibrate the model using Norwegian population panel data with nearly career-long earnings histories. After validating our model using a mandatory schooling reform, we gain insights into the underlying economic mechanisms that generate the effects of the policy and forecast the impacts of several policy alternatives.

We offer this handout to facilitate transdisciplinary collaboration in the future development of EKW models. We first describe their economic framework, mathematical formulation, and calibration procedure. We then turn to the seminal model outlined in Keane and Wolpin (1997) as an example and present our group's ensemble of research codes that allow for its specification, simulation, and calibration. Finally, we summarize our efforts drawing on research outside economics to address the computational challenges in applying EKW models and improve the reliability and interpretability of their results.

Throughout, we only offer a limited number of seminal references and textbooks that invite further study. We introduce acronyms and symbols as needed, and our notation draws on the reviews by Aguirregabiria and Mira (2010), Arcidiacono and Ellickson (2011), and Puterman (1994).

# 2 Setup

We now present the basic setup of the EKW models. We first describe the economic framework, then turn to its mathematical formulation, and finally outline the calibration procedure.

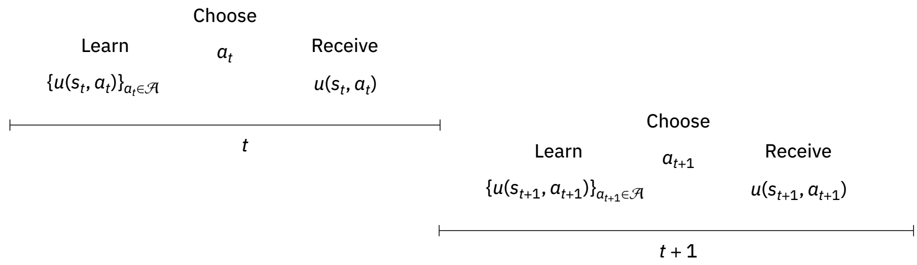



**Figure 1.** Timing of events

## 2.1 Economic framework

EKW models describe sequential decision-making under uncertainty (Gilboa, 2009; Machina & Viscusi, 2014). At time $t = 1, \ldots, T$ each individual observes the state of the economic 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 the economy 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 $\pi = (a_1^\pi(s_1), \ldots, a_T^\pi(s_T))$ provides the individual with instructions for choosing an action in any possible future state. It is a sequence of decision rules $a_t^\pi(s_t)$ that specify the action at a particular time $t$ for any possible state $s_t$ under $\pi$. 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. Individuals have rational expectations (Muth, 1961) so their subjective beliefs about the future agree with the objective transition probabilities of the model.

Figure 1 depicts the timing of events in the model for two generic periods. At the beginning of period $t$, an individual fully learns about the immediate utility of each alternative, 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 face uncertainty and they seek to maximize the expected total discounted utilities. An exponential discount factor $0 < \delta < 1$ parameterizes their time preference and captures a taste for immediate over future utilities.

Equation (1) provides the formal representation of the individual's objective. Given an initial state $s_1$, individuals implement the policy $\pi$ from the set of all possible policies $\Pi$ that maximizes the expected total discounted utilities over all $T$ decision periods given the information $\mathcal{I}_1$ available in the first period:

$$\max_{\pi \in \Pi} \mathrm{E}^{\pi}_{s_1} \left[ \sum_{t=1}^{T} \delta^{t-1} u_t(s_t, a^{\pi}_t(s_t)) \,\middle|\, \mathcal{I}_1 \right]. \tag{1}$$

The superscript of the expectation emphasizes that each policy $\pi$ induces a different probability distribution over the sequences of utilities.

## 2.2 Mathematical formulation

EKW models are set up as a standard Markov decision process (MDP) (Puterman, 1994; White, 1993). When making sequential decisions under uncertainty, the task is to determine the optimal policy $\pi^*$ with the largest expected total discounted utilities $v^{\pi^*}_1(s_1)$ as formalized in equation (1). In principle, this requires evaluating the performance of all policies based on all possible sequences of utilities, each weighted by the probability with which they occur. Fortunately, however, the multistage problem can be solved by a sequence of simpler inductively defined single-stage problems.[1]

The value function $v^{\pi}_t(s_t)$ captures the expected total discounted utilities under policy $\pi$ from period $t$ onwards for an individual experiencing state $s_t$:

$$v^{\pi}_t(s_t) = \mathrm{E}^{\pi}_{s_t} \left[ \sum_{j=0}^{T-t} \delta^j \, u_{t+j}(s_{t+j}, a^{\pi}_{t+j}(s_{t+j})) \,\middle|\, \mathcal{I}_t \right].$$

Then we can determine $v^{\pi}_1(s_1)$ for any policy by recursively evaluating equation (2):

$$v^{\pi}_t(s_t) = u_t(s_t, a^{\pi}_t(s_t)) + \delta \; \mathrm{E}^{\pi}_{s_t} \left[ v^{\pi}_{t+1}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right]. \tag{2}$$

Equation (2) expresses the total value $v^{\pi}_t(s_t)$ of adopting policy $\pi$ going forward as the sum of its immediate utility and all expected discounted future utilities.

The principle of optimality (Bellman, 1954) allows to construct $\pi^*$ by solving the optimality equations (3) for all $s$ and $t$ recursively:

$$v^{\pi^*}_t(s_t) = \max_{a_t \in A} \left\{ u_t(s_t, a_t) + \delta \; \mathrm{E}^{\pi^*}_{s_t} \left[ v^{\pi^*}_{t+1}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \right\}. \tag{3}$$

The optimal value function $v^{\pi^*}_t$ is the sum of the expected discounted utilities in $t$ over the remaining time horizon assuming the optimal policy is implemented going forward. The optimal action is choosing the alternative with the highest total value:

$$a^{\pi^*}_t(s_t) = \arg\max_{a_t \in A} \left\{ u_t(s_t, a_t) + \delta \; \mathrm{E}^{\pi^*}_{s_t} \left[ v^{\pi^*}_{t+1}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \right\}.$$

1. Optimal decisions in an MDP are a deterministic function of the current state $s$ only, i.e., an optimal decision rule is always deterministic and Markovian. We restrict our notation to this special case right from the beginning.

**Algorithm 1.** Backward induction procedure

**for** $t = T, \ldots, 1$ **do**
  **if** t == T **then**
    $$v_T^{\pi^*}(s_T) = \max_{a_T \in A} \Big\{ u_T(s_T, a_T) \Big\} \qquad \forall\, s_T \in S$$
  **else**
    Compute $v_t^{\pi^*}(s_t)$ for each $s_t \in S$ by
    $$v_t^{\pi^*}(s_t) = \max_{a_t \in A} \Big\{ u_t(s_t, a_t) + \delta\ \mathrm{E}_{s_t}^{\pi}\left[ v_{t+1}^{\pi^*}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \Big\}$$
    and set
    $$a_t^{\pi^*}(s_t) = \arg\max_{a_t \in A} \Big\{ u_t(s_t, a_t) + \delta\ \mathrm{E}_{s_t}^{\pi}\left[ v_{t+1}^{\pi^*}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \Big\}.$$
  **end if**
**end for**

Algorithm 1 allows to solve the MDP by a simple backward induction procedure. In the final period $T$, there is no future to take into account, and the optimal action is choosing the alternative with the highest immediate utilities in each state. With the decision rule for the final period at hand, the other optimal decisions can be determined recursively following equation (3) as the calculation of their expected future utilities is straightforward given the relevant transition probabilities.

## 2.3 Calibration procedure

EKW models are calibrated to data on observed individual decisions and experiences under the hypothesis that the individual behaves according to the model. The goal is to back out information on utility functions, preference parameters, and transition probabilities. This requires the full parameterization $\theta$ of the model.

Economists have access to information 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 $s_{it} = (\bar{s}_{it}, \epsilon_{it})$. At time $t$, the economist and individual both observe $\bar{s}_{it}$ while $\epsilon_{it}$ is only observed by the individual. In summary, the data $\mathcal{D}$ has the following structure:

$$\mathcal{D} = \{a_{it}, \bar{s}_{it}, \bar{u}_{it} : i = 1, \ldots, N; t = 1, \ldots, T_i\},$$

where $T_i$ is the number of observations for which we observe individual $i$.

Numerous calibration procedures for different settings exist (Davidson & MacKinnon, 2003; Gourieroux & Monfort, 1996). We briefly outline likelihood-based and simulation-based calibration. Independent of the calibration criterion, it is necessary to solve for the optimal policy $\pi^*$ at each candidate parameterization of the model.

Likelihood-based calibration seeks to find the parameterization $\hat{\theta}$ that maximizes the likelihood function $\mathcal{L}(\theta \mid \mathcal{D})$, i.e. the probability of observing the given data as a function of $\theta$. As we only observe a subset $\bar{s}_t$ of the state, we can determine the probability $p_{it}(a_{it}, \bar{u}_{it} \mid \bar{s}_{it}, \theta)$ of individual $i$ at time $t$ in $\bar{s}_{it}$ choosing $a_{it}$ and receiving $u_{it}$ given parametric assumptions about the distribution of $\epsilon_{it}$. The objective function takes the following form:

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

In simulation-based calibration, our goal is to find the parameterization $\hat{\theta}$ that yields a simulated data set from the model that closest resembles the observed data. More precisely, the goal is often to minimize the weighted squared distance between a set of moments $M_D$ computed on the observed data and the same set of moments computed on the simulated data $M_S(\theta)$. The objective function takes the following form:

$$\hat{\theta} \equiv \arg\min_{\theta \in \Theta} (M_D - M_S(\theta))' W (M_D - M_S(\theta)).$$

## 3 Example

We now present an exemplifying analysis of a canonical EKW model on human capital investment. The model was initially studied in Keane and Wolpin (1997) to explore the career decisions of young men about their schooling, work, and occupational choice. We first outline the basic setup of the model, provide some descriptive statistics of the empirical data used for its calibration, and then explore selected economic insights.

### 3.1 Basic setup

We follow individuals over their working life from young adulthood at age 16 to retirement at age 65 where the decision period $t = 16, \dots, 65$ is a school year. Figure 2illustrates 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 ($a = 5$).

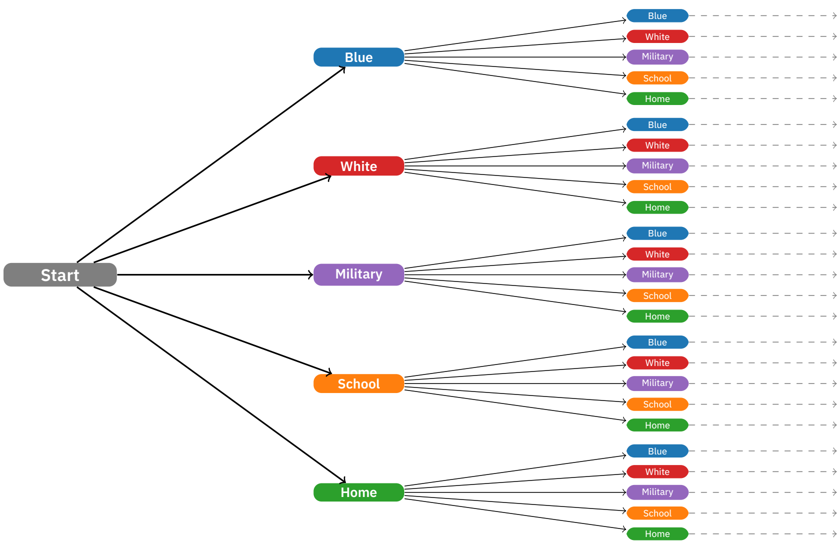



**Figure 2.** Decision tree

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

The immediate utility $u(\cdot)$ 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 $\boldsymbol{k}_t = \left(k_{a,t}\right)_{a\in\{1,2,3\}}$, years of completed schooling $h_t$, and alternative-specific skill endowment $\boldsymbol{e}$. The immediate utilities are influenced by last-period choices $a_{t-1}$ and alternative-specific productivity shocks $\boldsymbol{\epsilon}_t = \left(\epsilon_{a,t}\right)_{a\in\mathcal{A}}$ as well. Their general form is given by:

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

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

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

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

<<<<<<< HEAD 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 per-period utility in the blue-collar occupation as an example.[2] ======= 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 per-period utility in the blue-collar occupation as an example.[3]
>>>>>>> master

Equation (4) shows the parameterization of the non-pecuniary utility from working in a blue-collar occupation:

$$\begin{aligned}\zeta_1(\boldsymbol{k}_t, h_t, a_{t-1}) = \alpha_1 &+ c_{1,1} \cdot 1[a_{t-1} \neq 1] + c_{1,2} \cdot 1[k_{1,t} = 0] \\ &+ \vartheta_1 \cdot 1[h_t \geq 12] + \vartheta_2 \cdot 1[h_t \geq 16] + \vartheta_3 \cdot 1[k_{3,t} = 1].\end{aligned} \tag{4}$$

It includes job amenities $\alpha_1$ and mobility and search costs $(c_{1,1}, c_{1,2})$ that capture the extra effort for individuals who only recently started working in a blue-collar occupation. Additional components depend on whether an individual has a high school $\vartheta_1$ or college $\vartheta_2$ degree. There is a detrimental impact of leaving the military after a single year $\vartheta_3$.

The wage component $w_1(\cdot)$ is given by the product of the market-equilibrium rental price $r_1$ and an occupation-specific skill level $x_1(\cdot)$. The latter is determined by the overall level of human capital. This specification leads to a standard logarithmic wage equation in which the constant term is the skill rental price $\ln(r_1)$ and wages follow a log-normal distribution.

The occupation-specific skill level $x_1(\cdot)$ is determined by a skill production function, which includes a deterministic component $\Gamma_1(\cdot)$ and a multiplicative stochastic productivity shock $\epsilon_{1,t}$:

$$x_1(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,1}, \epsilon_{1,t}) = \exp\left(\Gamma_1(\boldsymbol{k}_t, h_t, t, a_{t-1}, e_{j,1}) \cdot \epsilon_{1,t}\right).$$

Equation (5) shows the parameterization of the deterministic component of the skill production function:

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

<<<<<<< HEAD ======= >>>>>>> master There are several notable features. Skills increase with schooling $\beta_{1,1}$ and blue-collar work experience $(\gamma_{1,1}, \gamma_{1,2})$. There

2. All additional details are available in Appendix ??.
3. All additional details are available in Appendix ??.

are so-called sheep-skin effects 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.

### 3.2 Empirical data

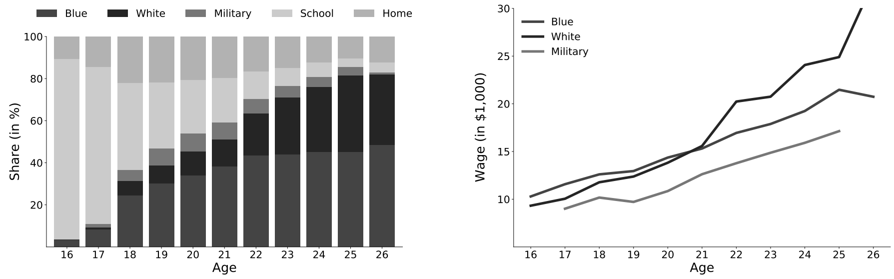
We analyze the original dataset used by Keane and Wolpin (1997) and thus only provide a brief description here.[4] The authors construct their sample based on the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of young men and women living in the United States in 1979 and born between 1957 and 1964. Individuals were followed from 1979 onwards and repeatedly interviewed about their 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, missing data, and the short observation period. Overall, the final sample consists of 12,359 person-period observations.

Figure 3 summarizes our information about choices and wages by age. We show the distribution of choices on the left, and report average wages on the right. Initially, roughly 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 occu-

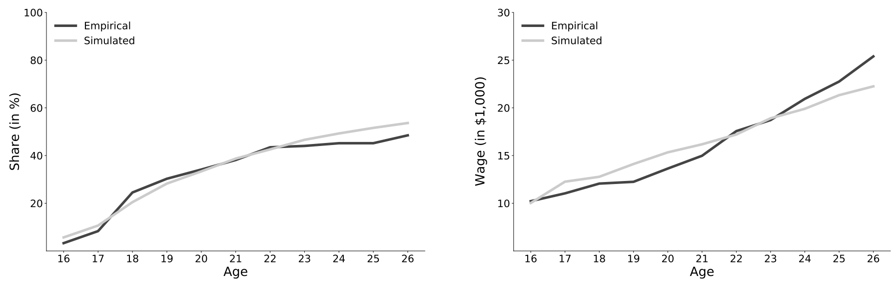
4. We provide additional details in Appendix **??**.




**(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 3.** Data overview




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

**Notes:** We simulate a sample of 1,000 individuals using the calibrated model.

**Figure 4.** Model fit

pation 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 fit the model to the empirical data using maximum likelihood calibration. Figure 4 shows the overall agreement between the empirical data and a dataset simulated using the calibrated parameters within the support of the data. On the left, we show the choice probability of working in a blue-collar occupation, while we plot the average wage across all occupations on the right.

Overall, the values of the calibrated parameters of the model are in broad agreement with the relevant literature. For example, individuals discount future utilities by 6% per year, and wages increase by about 7% with each additional year of schooling.

### 3.3 Economic insights

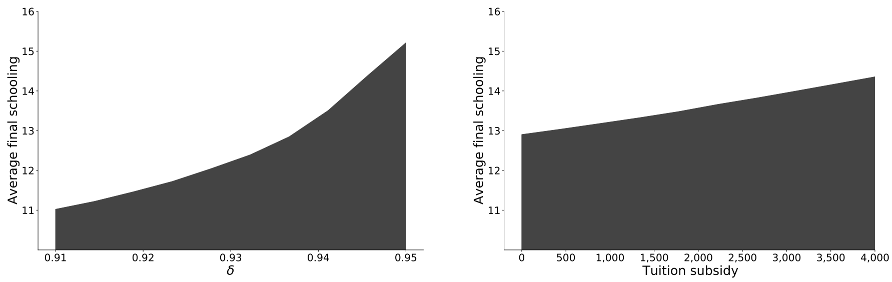
Figure 5 illustrates the ability of the model to quantify the impact of economic mechanisms and to forecast the effect of public policies. On the left, we vary the discount factor capturing time preferences between 0.91 and 0.95 while we introduce a tuition subsidy of up to $4,000 on the right. In both cases, we are interested in the changes to average final schooling.

Increases in the discount factor and the tuition subsidy both result in higher average final schooling. However, they do so for very different reasons. While individuals emphasize the future benefits of their schooling investment in the former, they react to a reduction of its immediate cost in the latter.

## 4 Pipeline

We are actively developing an ensemble of research codes that provide an analysis pipeline for EKW models. Among them are `respy` and `estimagic`. The former allows for the flexible specification and simulation of EKW models while the latter provides the means for their calibration. We briefly showcase the typical workflow of using both packages in our research.

Figure 6 illustrates a typical workflow. Initially, the user provides the empirical data, the parameterization of the model, and other options to `respy`. All together define the structure of the model, and we can construct the functionality for the simulation of data and the evaluation of the criterion function. `estimagic` allows calibrating the model to the empirical data. The results from




**(a)** Time preference

**(b)** Tuition subsidy

**Notes:** We simulate a sample of 1,000 individuals using the calibrated model.

**Figure 5.** Economic mechanism and policy forecast

```python
import respy as rp
from estimagic import maximize

# obtain model input
params, options, df = rp.get_example_model("kw_97_extended_respy")

# process model specification
log_like = rp.get_log_like_func(params, options, df)
simulate = rp.get_simulate_func(params, options)

# perform calibration
results, params_rslt = maximize(log_like, params, "nlopt_bobyqa")

# conduct analysis
df_rslt = simulate(params_rslt)
```

**Figure 6.** Typical workflow

| | | value | name |
|---|---|---|---|
| category | name | | |
| delta | delta | 9.370735e-01 | delta_delta |
| wage_white_collar | constant | 8.741888e+00 | wage_white_collar_constant |
| | exp_school | 6.548940e-02 | wage_white_collar_exp_school |
| | exp_white_collar | 1.763655e-02 | wage_white_collar_exp_white_collar |
| | exp_white_collar_square | -4.215936e-02 | wage_white_collar_exp_white_collar_square |
| | exp_blue_collar | 3.431936e-02 | wage_white_collar_exp_blue_collar |
| | exp_military | 1.406945e-02 | wage_white_collar_exp_military |
| | hs_graduate | -3.599855e-03 | wage_white_collar_hs_graduate |
| | co_graduate | 2.301313e-03 | wage_white_collar_co_graduate |
| | period | 9.577717e-03 | wage_white_collar_period |
| | is_minor | -1.509984e-01 | wage_white_collar_is_minor |

**(a)** Parameterization

| | value |
|---|---|
| estimation_draws | 200 |
| estimation_seed | 500 |
| estimation_tau | 500 |
| interpolation_points | -1 |
| n_periods | 50 |
| simulation_agents | 5000 |
| simulation_seed | 132 |
| solution_draws | 500 |
| solution_seed | 456 |
| monte_carlo_sequence | random |
| covariates | {'hs_graduate': 'exp_school >= 12', 'co_gradua... |

**(b)** Options

**Figure 7.** Model specification

the calibration steps are used to, for example, analyze the economic mechanisms underlying the observed behaviors.

Figure 7 shows the model specification files for Keane and Wolpin (1997). The file on the left sets the parameter values for the utility functions and the distribution of the unobservable state variables. On the right, we provide details on the construction of the observed state variables and numerous tuning parameters for the numerical solution of the model.

Figure 8 depicts the dashboard provided by `estimagic` to monitor the progress and parameter values of the calibration in real-time. This allows us to detect problems during calibration right away and facilitates the debugging process.

We adopt a modern software engineering workflow in the development of both packages and tutorials, source code, testing harness, as well as implementation details are available in their respective online documentations at https://respy.rtfd.io and https://estimagic.rtfd.io.

## 5 Improvements

The implementation of EKW models poses several computational challenges. Among them are numerical integration, global optimization, function approximation, and efficient parallelization. We now describe some of our efforts to align `respy` and `estimagic` with the state-of-the-art in computational methods. We have concluded our preparatory work and actively seek input from domain experts for further improvements and joint publication.

### 5.1 Numerical integration

The solution of EKW models requires the evaluation of millions of integrals to determine the future value of each action in each state. In Eisenhauer, Gabler, and Suchy (2020), we draw on the extensive literature on numerical integration (Davis & Rabinowitz, 2007; Gerstner & Griebel, 1998) to improve the precision and reliability of their solution. The current practice in economics is to implement a random Monte Carlo integration which introduces considerable numerical error and computational instabilities (Judd & Skrainka, 2011).

We consider the optimality equation in a generic time period $t$ to clarify the structure of the integral. Let $v_t^{\pi}(s_t, a_t)$ denote the action-specific value function of choosing action $a_t$ in state $s_t$ while continuing with the optimal policy going forward.

$$\begin{aligned} v_t^{\pi}(s_t, a_t) &= u_t(s_t, a_t) + \delta \; \mathrm{E}_{s_t}\left[ v_{t+1}^{\pi^*}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \\ &= u_t(s_t, a_t) + \delta \int_S v_{t+1}^{\pi^*}(s_{t+1}) \; \mathrm{d}p_t(a_t, s_t) \\ &= u_t(s_t, a_t) + \delta \int_S \max_{a_{t+1} \in A} \left\{ v_{t+1}^{\pi^*}(s_{t+1}, a_{t+1}) \right\} \mathrm{d}p_t(a_t, s_t). \end{aligned}$$

Let's consider an atemporal version of the typical integral from Keane and Wolpin (1997) as an example. As outlined earlier, individuals can choose

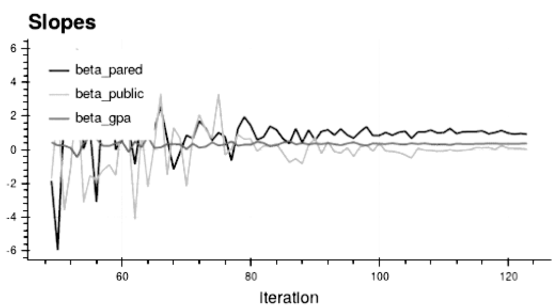



**Figure 8.** Dashboard

among five alternatives. Each of the alternative-specific utilities is, in part, determined by a stochastic continuous state variable $\boldsymbol{\varepsilon}$. The transition of all other state variables $x$ is deterministic. This results in a five-dimensional integral of the following form:

$$\int \max_{a \in A} \left\{ V^{\pi^*}(x, \boldsymbol{\varepsilon}, a) \right\} \phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\varepsilon}) \, \mathrm{d}\boldsymbol{\varepsilon} \quad \forall \, x \in X,$$

where $\boldsymbol{\varepsilon}$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$, and probability density function $\phi_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$.

## 5.2 Global optimization

The calibration of EKW models is challenging due to a large number of parameters and multiplicity of local minima. In Eisenhauer, Gabler, and Röhrl (2020), we draw on the literature on global optimization to assess and improve the reliability of the calibrations (Locatelli & Schoen, 2013; Nocedal & Wright, 2006).

We conduct a benchmarking exercise using Keane and Wolpin (1997) and Keane and Wolpin (1994) as a well-known and empirically grounded test case. Depending on the calibration procedure, particular challenges arise. For example, while likelihood-based calibration requires smoothing of the choice probabilities, simulation-based calibration involves the optimization of a noisy function. We provide guidelines for selecting the appropriate algorithm in each setting and showcase diagnostics to assess the reliability of the calibration results.

## References


Adda, J., Dustmann, C., & Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, *125*(2), 293–337. [3]

Aguirregabiria, V., & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, *156*(1), 38–67. [3]

Arcidiacono, P., & Ellickson, P. B. (2011). Practical methods for estimation of dynamic discrete choice models. *Annual Review of Economics*, *3*(1), 363–394. [3]

Becker, G. S. (1964). *Human capital*. Columbia University Press. [3]

Bellman, R. E. (1954). The theory of dynamic programming. *Bulletin of the American Mathematical Society*, *60*(6), 503–515. [5]

Bhuller, M., Eisenhauer, P., & Mendel, M. (2018). The option value of schooling. *Unpublished Working Paper*. [3]

Blundell, R., Costa Dias, M., Meghir, C., & Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, *84*(5), 1705–1753. [3]

Davidson, R., & MacKinnon, J. G. (2003). *Econometric theory and methods*. Oxford University Press. [6]

Davis, P. J., & Rabinowitz, P. (2007). *Methods of numerical integration*. Dover Books on Mathematics. [14]

Eisenhauer, P., Gabler, J., & Röhrl, K. (2020). Global optmization in structural econometric models. *Unpublished Working Paper*. [15]

Eisenhauer, P., Gabler, J., & Suchy, R. (2020). Advanced numerical integration in structural econometric models. *Unpublished Working Paper*. [14]

Gerstner, T., & Griebel, M. (1998). Numerical integration using sparse grids. *Numerical algorithms*, *18*(3-4), 209. [14]

Gilboa, I. (2009). *Theory of decision under uncertainty*. Cambridge University Press. [4]

Gourieroux, C., & Monfort, A. (1996). *Simulation-based econometrics*. Oxford University Press. [6]

Judd, K. L., & Skrainka, B. S. (2011). High performance quadrature rules: How numerical integration affects a popular model of product differentiation. *CEMMAP Working Paper (CWP03/11)*. [14]

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522. [1, 3, 7, 10, 13–15]

Keane, M. P., & Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Review of Economics and Statistics*, *76*(4), 648–672. [15]

Locatelli, M., & Schoen, F. (2013). *Global optimization*. SIAM. [15]

Machina, M. J., & Viscusi, K. (Eds.). (2014). *Handbook of the economics of risk and uncertainty* (Vol. 1). North Holland Publishing Company. [4]

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*(3), 315–335. [4]

Nocedal, J., & Wright, S. J. (2006). *Numerical optimization*. Springer. [15]

OECD. (2001). *The well-being of nations: The role of human and social capital*. OECD Publication Service. [3]

Puterman, M. L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons. [3, 5]

White, D. J. (1993). *Markov decision processes*. John Wiley & Sons. [5]

Wolpin, K. I. (2013). *The limits to inference without theory*. MIT University Press. [3]

## Contents

# Appendix A Computational implementation

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 $\boldsymbol{e} = \left(e_{j,a}\right)_{\mathcal{J}\times\mathcal{A}}$.

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 $\boldsymbol{k}_t = \left(k_{a,t}\right)_{a\in\{1,2,3\}}$, 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 $\boldsymbol{\varepsilon}_t = \left(\epsilon_{a,t}\right)_{a\in\mathcal{A}}$ 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:

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

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

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.

## A.1 Non-pecuniary utility

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

**Blue-collar.** Equation (A.1) shows the parameterization of the non-pecuniary utility from working in a blue-collar occupation:

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

A constant $\alpha_1$ 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 $\vartheta_1$ and a college $\vartheta_2$ degree. Additionally, there is a detrimental effect of leaving the military early after one year $\vartheta_3$.

**White-collar.** The non-pecuniary utility from working in a white-collar occupation is specified analogously. Equation (A.2) shows its parameterization:

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

**Military.** Equation (A.3) shows the parameterization of the non-pecuniary utility from working in the military:

$$\zeta_3(k_{3.t}, h_t) = c_{3,2} \cdot 1[k_{3,t} = 0] + \vartheta_1 \cdot 1[h_t \geq 12] + \vartheta_2 \cdot 1[h_t \geq 16]. \tag{A.3}$$

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 $\vartheta_1$ and college $\vartheta_2$.

**School.** Equation (A.4) shows the parameterization of the non-pecuniary utility from schooling:

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

There is a direct cost of attending school such as tuition for continuing education after high school $\beta_{tc_1}$ and college $\beta_{tc_2}$. 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 (A.5) shows the parameterization of the non-pecuniary utility from staying at home:

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

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 $\vartheta_3$ influences the immediate utility.

## A.2 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 $\epsilon_{a,t}$:

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

**Blue-collar.** Equation (A.6) shows the parameterization of the deterministic component of the skill production function:

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

«««< HEAD There are several notable features. The first part of the skill production function is motivated by Mincer (1958, 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 (Jaeger & Page, 1996; Spence, 1973) 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. ======= There are several notable features. The first part of the skill production function is motivated by Mincer (1958, 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 & Solon, 1987; Jaeger & 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. »»»> master

**White-collar.** The wage component from working in a white-collar occupation is specified analogously. Equation (A.7) shows the parameterization of the deterministic component of the skill production function:

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

**Military.** Equation (A.8) shows the parameterization of the deterministic component of the skill production function:

$$
\begin{aligned}
\Gamma_3(k_{3,t}, h_t, t, e_{j,3}) = e_{j,3} &+ \beta_{3,1} \cdot h_t \\
&+ \gamma_{3,1} \cdot k_{3,t} + \gamma_{3,2} \cdot (k_{3,t})^2 + \gamma_{3,3} \cdot 1[k_{3,t} > 0] \\
&+ \gamma_{3,4} \cdot t + \gamma_{3,5} \cdot 1[t < 18].
\end{aligned} \tag{A.8}
$$

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.3 Overview of parameters

**Table A.1.** Overview of parameters in the Keane and Wolpin (1997) extended model.

| Parameter | Description |
|---|---|
| Preference and type-specific parameters | |
| $\delta$ | discount factor |
| $e_{j,a}$ | initial endowment of type $j$ in alternative $a$ specific skills |
| Common parameters immediate utility | |
| $\alpha_a$ | return non-wage working conditions |
| $\vartheta_1$ | non-pecuniary premium of finishing high-school |
| $\vartheta_2$ | non-pecuniary premium finishing college |
| $\vartheta_3$ | non-pecuniary premium of leaving military early |
| Schooling-related parameters | |
| $\beta_{a,1}$ | return additional year of completed schooling |
| $\beta_{a,2}$ | skill premium high-school graduate |
| $\beta_{a,3}$ | skill premium college graduate |
| $\beta_{tc_1}$ | tuition cost high-school |
| $\beta_{tc_2}$ | tuition cost college |
| $\beta_{rc_1}$ | re-entry cost high-school |
| $\beta_{rc_2}$ | re-entry cost college |
| $\beta_{5,2}$ | skill premium high-school graduate |
| $\beta_{5,3}$ | skill premium college graduate |
| Experience-related parameters | |
| $\gamma_{a,1}$ | return same-sector experience |
| $\gamma_{a,2}$ | return squared same-sector experience |
| $\gamma_{a,3}$ | premium having worked in sector before |
| $\gamma_{a,4}$ | return age effect |
| $\gamma_{a,5}$ | return age effect being a minor |
| $\gamma_{a,6}$ | premium remaining in same sector |
| $\gamma_{a,7}$ | return civilian cross-sector experience |
| $\gamma_{a,8}$ | return non-civilian sector experience |
| $\gamma_{3,1}$ | return same-sector experience |

continued on next page

continued from previous page

| | |
|---|---|
| $\gamma_{3,2}$ | return squared same-sector experience |
| $\gamma_{3,3}$ | premium having worked in sector before |
| $\gamma_{3,4}$ | return age effect |
| $\gamma_{3,5}$ | return age effect being a minor |
| $\gamma_{4,4}$ | return age effect |
| $\gamma_{4,5}$ | return age effect being a minor |
| $\gamma_{5,4}$ | return age between 17 and 21 |
| $\gamma_{5,5}$ | return age older than 20 |
| Mobility and search parameters | |
| $c_{a,1}$ | premium switching to occupation $a$ |
| $c_{a,2}$ | premium for working first time in occupation $a$ |
| $c_{3,2}$ | premium for serving first time in military |
| Error correlation | |
| $\sigma_{a,a}$ | standard deviation of shock in alternative $a$ |
| $\sigma_{i,j}$ | correlation between shocks of alternative $a = i$ and $a = j$ with $i \neq j$ |

**Note:** The listed parameters are represented as an overview. The immediate utilities for the alternatives do not necessarily include all of them.

## Appendix B Empirical data

We use the same data as in Keane and Wolpin (1997). Please see https://bit.ly/ekw-data for details.

## References

Altonji, J. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, *11*(1), 48–83. [4]

Hungerford, T., & Solon, G. (1987). Sheepskin effects in the returns to education. *Review of Economics and Statistics*, *69*(1), 175–177. [5]

Jaeger, D. A., & Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, *78*(4), 733–740. [5]

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522. [2, 6, 7, 9]

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, *66*(4), 281–302. [5]

Mincer, J. (1974). *Schooling, experience and earnings*. National Bureau of Economic Research. [5]

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. [5]