# Eckstein-Keane-Wolpin models*

OpenSourceEconomics

**Abstract**

This notes provides some background material to understand the class of Eckstein-Keane-Wolpin (EKW) models. EKW models are commonly used in the context of labor economics. These can be solved, simulated, and estimated using the `respy` package.

**JEL Codes:** J24, D81, C44

**Keywords:** life cycle model, human capital, risk, Markov decision process

*Department of Economics, University of Bonn. Email: peisenha@uni-bonn.de. I gratefully acknowledge support by the AXA Research Fund. I thank Jörg Stoye, Janoś Gabler, and Rafael Suchy for their comments and suggested improvements. I thank the editor and two anonymous referees for their constructive comments that greatly improved the manuscript. I am grateful to the Social Sciences Computing Service at the University of Chicago for the permission to use their computational resources.

# Contents

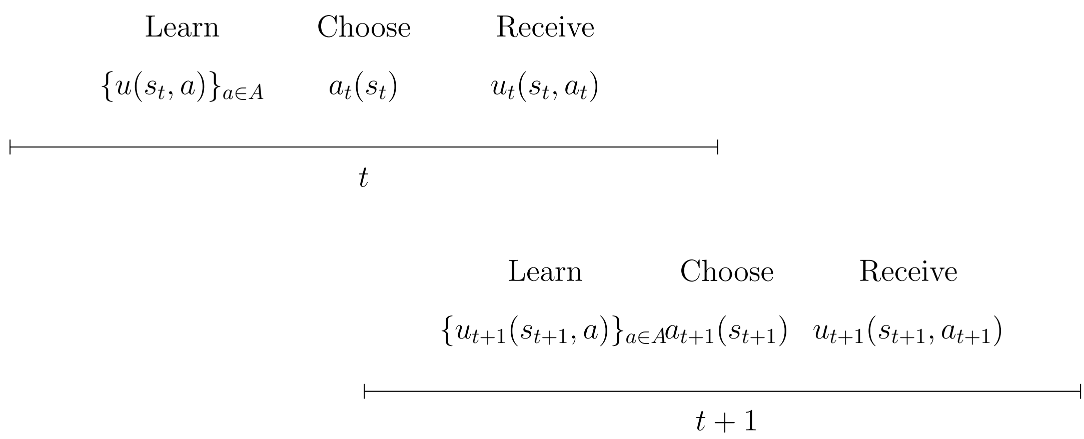
# 1. Introduction

This handout provides an overview regarding our efforts in the implementation ...
We now present the economic, mathematical, and computational model for the class of Eckstein-Keane-Wolpin (EKW) model. We will start with a discussion of the basic economic environment and then turn to the corresponding mathematical model of standard Markov decision process (MPD). We continue with a brief description of the estimation procedure before concluding this section with a discussion of some computational challenges when working with these models. Throughout, we will introduce acronyms and symbols as needed, but a full list of both is provided in Appendix A. The notation draws form the related work by Puterman (1994), Aguirregabiria & Mira (2010), and Arcidiacono & Ellickson (2011).

**Computational challenges** The implementation and analysis of this class of models entails several computational challenges. Among them integration of a high-dimensional non-differentiable function, large-scale global optimization of a noisy and non-smooth criterion function, function approximation, and parallelization strategies. We briefly outline each of them.

**Stucture** The remainder of this handout is structured as follows. We first present the economic model with a discussion of the basic economic environment. In Section 3 we present the corresponding mathematical model of a standard finite-horizon discrete Markov decision process (MPD) and outline its solution approach. When then outline the estimation step in Section 4. With this overview at hand, we discuss selected computational challenges in Section 5. This note concludes with an example model in Section 6.

**Figure 1:** Timing of events




## 2. Economic motivation

**Basic setup** 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 $u(s_t, a_t)$ 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. 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.

**Decision rule** A policy $\pi \equiv (a_1^\pi(s), \ldots, a_T^\pi(s))$ provides the individual with a prescription for choosing an action in any possible future state, where $a_t^\pi(s)$ specifies the action at a particular time $t$ for any possible state under $\pi$. It is a sequence of decision rules and its implementation generates a sequence of utilities. The evolution of states over time is at least partly unknown as future utilities depend on, for example, shocks to preferences. Individuals use models about their economic environment to inform their beliefs about the future. For a given model, individuals thus face risk as each induces a unique objective transition probability distribution $p_t(s_t, a_t)$ for the evolution of state $s_t$ to $s_{t+1}$ that depends on the action $a_t$.

**Timing of events** Figure 1 depicts the timing of events in the model for two generic time periods. At the beginning of time $t$ 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 $t + 1$.

**Decision theory** Individuals make their decisions facing risk and have rational expectations (Muth, 1961; Lucas, 1972) as their model about the future also turns out to be true. In this case, there is a consensus that rational choices are expressed by expected utility preferences

(Bernoulli, 1738; von Neumann & Morgenstern, 1944, 1947).

**Formalization** Individuals maximize their expected total discounted utility (Samuelson, 1937; Koopmans, 1960). A constant discount factor ensures dynamic consistency of preferences as the individual's future actions agree with the planned-for contingencies. Beliefs are updated according to Bayes's rule.[1]

Equation (1) provides the formal representation of the individual's objective. Given an initial state $s_1$, individuals seek to implement the optimal policy $\pi^*$ from the set of all possible policies $\Pi$ that maximizes the expected total discounted utility over all $T$ decision periods given the information available at the time $\mathcal{I}_1$.

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

The exponential discount factor $0 < \delta < 1$ captures a preference for immediate over future utility. The superscript of the expectation emphasizes that each policy $\pi$ induces a different unique probability distribution over the sequences of utilities.

[1] See Frederick et al. (2002) for a critical review of the literature on time discounting and time preference. Fang & Silverman (2009), Fang & Yang (2015), and Chan (2017) are examples of hyperbolic discounting and thus potentially time-inconsistent preferences in settings similar to the one discussed here.

# 3. Mathematical formulation

EKW models are set up as a standard MDP.[2] When making sequential decisions, the task is to determine the optimal policy $\pi^*$ with the largest expected total discounted utility $v_1^{\pi^*}$ as formalized in equation (1). In principle, this requires to evaluate the performance of all policies based on all possible sequences of utilities and the probability that each occurs. Fortunately, however, the multistage problem can be solved by a sequence of simpler inductively defined single-stage problems. Note that in slight abuse of notation $s_{t+1}$ is a random variable given the information available at $\mathcal{I}_t$.

Let $v_t^\pi(s)$ denote the expected total discounted utility under $\pi$ from period $t$ onwards:

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

Then $v_1^\pi(s_t)$ can be determined for any policy by recursively evaluating equation (2):

$$v_t^\pi(s) = u(s_t, a_t^\pi(s_t)) + \delta \, \mathrm{E}_{s_t} \left[ v_{t+1}^\pi(s_{t+1}) \,\middle|\, \mathcal{I}_t \right]. \tag{2}$$

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

The principle of optimality (Bellman, 1957; Puterman, 1994) allows to construct the optimal policy $\pi^*$ by solving the optimality equations for all $s$ and $t$ in equation (3) recursively:

$$v_t^{\pi^*}(s_t) = \max_{a \in A} \left\{ u(s_t, a) + \delta \, \mathrm{E}_{s_t} \left[ v_{t+1}^{\pi^*}(s_{t+1}) \,\middle|\, \mathcal{I}_t \right] \right\} \tag{3}$$

The value function $v_t^{\pi^*}$ is the expected discounted utility in $t$ over the remaining time horizon assuming the optimal policy is implemented going forward.

The optimal decision is simply the alternative with the highest value:

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

**Solution approach** 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 so the optimal decision is simply to choose the alternative with the highest immediate utility in each state.

[2] See Puterman (1994) and White (1993) for a textbook introduction to standard Markov decision processes and Rust (1994) for a review of its use in structural estimation.

With the results for the final period at hand, the other optimal decisions can be determined recursively as the calculation of their expected future utility is straightforward given the relevant transition probability distribution.

---

**Algorithm 1** Backward Induction Algorithm for MDP

---

**for** $t = T, \ldots, 1$ **do**

  **if** t == T **then**

    $$v_T^{\pi^*}(s_T) = \max_{a \in A} \Big\{ u(s_T, a) \Big\} \qquad \forall\, s_T \in S$$

  **else**

    Compute $v_t^{\pi^*}(s_t)$ for each $s_t \in S$ by

    $$v_t^{\pi^*}(s_t) = \max_{a \in A} \Big\{ u(s_t, a) + \delta\, \mathrm{E}_{s_t} \left[ v_{t+1}^{\pi^*}(s_{t+1}) \middle|\, \mathcal{I}_t \right] \Big\}$$

    and set

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

  **end if**

**end for**

---

## 4. Estimation procedures

**Available information** The econometrician has access to panel data for $N$ individuals. For every observation $(i,t)$ in the panel data set, the researcher observes action $a_{it}$ and a subvector $x_{it}$ of the state vector. Therefore, from an econometricians point of view, we need to distinguish between two types of state variables $s_{it} = (x_{it}, \epsilon_{it})$. Variables $x_{it}$ that are observed by the econometrician and the individual $i$ at time $t$ and those that are only observed by the individual $\epsilon_{it}$. In addition, also some realizations of the rewards $r_{it} = r(x_{it}, \epsilon_{it}, a_{it})$.

$$\mathcal{D} = \{a_{it}, x_{it}, r_{it} : i = 1, 2, \ldots, N; t = 1, \ldots, T_i\}\},$$

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

**Parameterization** The models need to be parameterized functional forms and the distribution of unobservables. EKW models are calibrated to obtain information on structural parameters of preferences and the transition probabilities that reproduce key economic patterns of interest in observed sources such as administrative data sets and the like. This allows to assess the quantitative importance of competing economic mechanisms and then predict the effects of public policies. We collect all parameters of the model in $\theta$.

**Procedures** We briefly outline maximum likelihood estimation (Fisher, 1922) and the method of simulated moments (McFadden, 1989). Whatever the estimation criterion, in order to evaluate it for a particular value of $\theta$ it is necessary to know the optimal decision rules $d^{\pi^*}$. Therefore at each trial value of $\theta$ the dynamic programming model needs to be solved. More detailed information is available in the excellent textbooks by Davidson & MacKinnon (2003) and Gourieroux & Monfort (1996).

**Likelihood-based estimation** The individual chooses the alternative with the highest total value $a_t^*(s_t)$ which is determined by the complete state space and rewards are also determined by $s$. However, the econometrician only observes the subset $x$ and thus can only determine the probability $p_{it}(a_{it}, r_{it} \mid x_{it}, \theta)$ of individual $i$ at time $t$ choosing $d_{it}$ and receiving $r_{it}$ given $x_{it}$.

$$\mathcal{L}(\theta \mid \mathcal{D}) = \prod_{i=1}^{N} \prod_{t=1}^{T_i} p_{it}(a_{it}, r_{it} \mid x_{it}, \theta)$$

The goal of likelihood-based estimation is to find the value of the model parameters $\theta$ that maximize the likelihood function:

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\max}\, \mathcal{L}(\theta \mid \mathcal{D})$$

**Simulation-based estimation** ...

# 5. Computational challenges

The class of EKW models has several key characteristics that distinguish it from other classes of models but entail several computational complications. We will first discuss its key features and then describe some of the computational challenges.

**Structure of integral** To clarify the structure of the integral determining the future value of a state, it is useful to consider the optimality equation in the second to last period. This allows to focus on action-specific rewards instead of future values. Let $v^{\pi}_{T-1}(s_t, a_t)$ denote the action-specific value function of choosing action $a$ in state $s$ while continuing with the optimal policy going forward. but sticking to the optimal policy $\pi^*$ going forward.

$$
\begin{aligned}
v^{\pi}_t(s_t, a_t) &= u(s_t, a_t) + \delta\, \mathrm{E}_{s_t}\left[v^{\pi^*}_{t+1}(s_{t+1})\right] &\quad (4)\\
&= u(s_t, a_t) + \delta \int_S v^{\pi^*}_{t+1}(s_{t+1})\, \mathrm{d}p_t(a_t, s_t) &\quad (5)\\
&= u(s_t, a_t) + \delta\, \underbrace{\int_S \max_{a\in A}\left\{v^{\pi}_{t+1}(s_{t+1}, a)\right\} \mathrm{d}p_t(a_t, s_t)}_{\mathcal{I}(a_{t+1})}. &\quad (6)
\end{aligned}
$$

The evaluation of such an integral is required millions of times during the backward induction procedure. The current practice is to implement a random Monte Carlo integration which introduces considerable numerical error and computational instabilities (Judd & Skrainka, 2011).

Let's consider an atemporal version of the typical integral from Keane & Wolpin (1997). In their model, individuals can choose among five alternatives. Each of the alternative-specific rewards is in part determined by a random continuous state variable that follows a normal distribution which happens to be unobserved. The transition of all observable state variables is deterministic. This results in a five-dimensional integral $\mathcal{I}(a')$ as the dimensionality is determined by the random state variables. The integral takes the following form:

$$
\mathcal{I}(a) = \int_{\epsilon} \max_{a\in A}\left\{v^{\pi}_{t+1}(x, \epsilon, a)\right\}\phi_{\mu,\Sigma}(\epsilon)\, \mathrm{d}\epsilon.
$$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_5) \sim \mathcal{N}(\mu, \Sigma)$ follows a multivariate normal distribution with mean $\mu \in \mathbb{R}^5$, covariance matrix $\Sigma \in \mathbb{R}^{5\times 5}$, and probability density function $\phi_{\mu,\Sigma}(\epsilon)$. A key features of this integral is the lack of general separability between the random and deterministic state variables. This makes the closed-form solutions impossible even under suitable distributional assumptions (?Rust, 1987).

**Global optimization** ...

- **Likelihood-based estimation** This approach requires smoothing of the choice probabilities.

$$p_t(d_{it} \mid x_{it}, \theta) = \int \mathbb{I}\left[\delta(x_{it}, \epsilon_{it}, \theta) = a_{it}\right] g(\epsilon)\, \mathrm{d}\epsilon$$

- **Simulation-based estimation** This approach requires the optimization of a noisy function.

**Function approximation** ...

**Parallelization** ...

# 6. Example

We now provide either Keane & Wolpin (1994) or Keane & Wolpin (1997) here. Given the purpose of this note as a high level overview, it is probably fine to just restrict to (Keane & Wolpin, 1994).

# 7. Extensions

We briefly discuss selected extensions to the baseline model that are our active areas of research:

## 7.1. Robust decision-making

Individuals face ubiquitous uncertainties when faced with important decisions. Policy makers vote for climate change mitigation efforts facing uncertainty about future costs and benefits (Barnett et al., 2019), while doctors decide on the timing of an organ transplant in light of uncertainty about future patient health (Kaufman et al., 2017). Economic models formalize the objectives, trade-offs, and uncertainties for such decisions. In these models, the treatment of uncertainty is often limited to risk as the model induces a unique probability distribution over sequences of possible futures. There is no role for ambiguity about the true model (Knight, 1921; Arrow, 1951) and thus no fear of model misspecification. However, limits to knowledge lead to considerable ambiguity about how the future unfolds (Hayek, 1975; Hansen, 2015).

This creates the need for robust decision rules that work well over a whole range of different models, instead of a decision rule that is optimal for one particular model. Optimal decision rules are designed without any fear of misspecification, using a single model to inform decisions. They thus perform very well if that model turns out to be true. However, their performance is very sensitive and deteriorates rapidly in light of model misspecification. Robust decision rules explicitly account for such a possibility and their performance is less affected.

Methods from distributionally robust optimization (Ben-Tal et al., 2009; Wiesemann et al., 2014; Rahimian & Mehrotra, 2019) and robust Markov decision processes (Iyengar, 2005; Nilim & El Ghaoui, 2005; Wiesemann et al., 2014) allow to construct decision rules that explicitly take potential model misspecification into account.

## 7.2. Uncertainty quantification

# A. Acronyms and Symbols

**Table 1:** List of Acronyms

| **Acronym** | **Meaning** |
|---|---|
| MDP | standard Markov decision process |

**Table 2:** List of Symbols

| **Symbol** | **Meaning** |
|---|---|
| $\mathbb{R}$ | set of real numbers |
| $\mathbb{I}[\,A\,]$ | indicator function that takes value one if event $A$ is true |
| | Economic Model |
| $t$ | decision period |
| $T$ | number of decision periods |
| $a \in \mathcal{A}$ | set of admissible actions |
| $s \in \mathcal{S}$ | set of possible states with generic state $s$ |
| $s_t$ | realization of state $s$ in period $t$ |
| $a_t^\pi(s)$ | decision rule that specifies an action for all states $s$ in period $t$ following $\pi$ |
| $a_t^\pi(s_t)$ | actual decision in period $t$ when in state $s_t$ following policy $\pi$ |

**Table 2:** List of Symbols

| Symbol | Meaning |
|---|---|
| $a_{it}$ | actual decision observed by individual $i$ at time $t$ |
| $p_t(s, a)$ | conditional probability distributions for $s_{t+1}$ when choosing action $a$ in state $s$ in period $t$ |
| $u(s, a)$ | utility when choosing action $a$ in state $s$ |
| $\delta$ | discount factor |
| $v_t^{\pi}$ | expected total discounted utility of adopting policy $\pi$ from period $t$ going forward |
| $\pi \in \Pi$ | set of all policies |
| Estimation procedure | |
| Computational Model | |
| $\epsilon_{at}$ | random shock to utility of alternative $a$ in period $t$ |
| $x_{jt}$ | number of periods worked in occupation $j$ by the beginning of period $t$ |
| $g_t$ | number of periods enrolled in school by the beginning of period $t$ |
| $\mathcal{N}_0$ | true multivariate normal distribution for random shocks |
| $\Sigma$ | covariance matrix of random shocks |
| $\upsilon$ | admissible realization of means for future labor market shocks |

**Table 2:** List of Symbols

| Symbol | Meaning |
|---|---|
| $\alpha_j$ | parameters for utility function when working in occupation $j$ |
| $\beta$ | parameters for utility function when enrolling in school |
| $\gamma$ | parameter for utility function when staying at home |

## References


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

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

Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, *19*(4), 404–437.

Barnett, M., Brock, W., & Hansen, L. P. (2019). Pricing uncertainty induced by climate change. *Working Paper*.

Bellman, R. E. (1957). *Dynamic programming.* Princeton, NJ: Princeton University Press.

Ben-Tal, A., El Ghaoui, L., & Nemirowski, A. (2009). *Robust optimization.* Princeton, NJ: Princeton University Press.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Commentarii academiae scientiarum imperialis Petropolitanae*, *5*, 175–192.

Chan, M. K. (2017). Welfare dependence and self-control: An empirical analysis. *Review of Economic Studies*, *84*(4), 1379–1423.

Davidson, R., & MacKinnon, J. G. (2003). *Econometric theory and methods.* New York: Oxford University Press.

Fang, H., & Silverman, D. (2009). Time-inconsistency and welfare program participation: Evidence from the NLSY. *International Economic Review*, *50*(4), 1043–1077.

Fang, H., & Yang, W. (2015). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *International Economic Review*, *56*(2), 565–596.

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.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351-401.

Gourieroux, C., & Monfort, A. (1996). *Simulation-based econometrics.* Oxford, United Kingdom: Oxford University Press.

Hansen, L. P. (2015). Uncertainty outside and inside economic models. *Journal of Political Economy*, *122*(5), 945–987.

Hayek, F. A. v. (1975). The pretence of knowledge. *The Swedish Journal of Economics*, *77*(4), 433–442.

Iyengar, G. N. (2005). Robust dynamic programming. *Mathematics of Operations Research*, *30*(2), 257–280.

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

Kaufman, D. L., Schaefer, A. J., & Roberts, M. S. (2017). Living-donor liver transplantation timing under ambiguous health state transition probabilities. *Working Paper*.

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.

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

Knight, F. H. (1921). *Risk, uncertainty and profit.* Boston, MA: Houghton Mifflin Harcourt.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, *28*(2), 287–309.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, *4*(2), 103-124.

McFadden, D. L. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, *57*(5), 995–1026.

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

Nilim, A., & El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, *53*(5), 780–798.

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

Rahimian, H., & Mehrotra, S. (2019). Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659 [math.OC]*.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, *55*(5), 999–1033.

Rust, J. (1994). Structural estimation of Markov decision processes. In R. Engle & D. McFadden (Eds.), *Handbook of Econometrics* (Vol. 4, pp. 3081–3143). Amsterdam, Netherlands: North-Holland.

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

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

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

Wiesemann, W., Kuhn, D., & Sim, M. (2014). Distributionally robust convex optimization. *Operations Research*, *62*(6), 1358–1376.