ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 104

Robust Decision-Making Under Risk and Ambiguity

Maximilian Blesch Philipp Eisenhauer

July 2021

www.econtribute.de

UNIVERSITÄT BONN Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Robust decision-making under risk and ambiguity*

Maximilian Blesch[1] and Philipp Eisenhauer[2]

[1]Berlin School of Economics

[2]University of Bonn

July 6, 2021

Economists often estimate a subset of their model parameters outside the model and let the decision-makers inside the model treat these point estimates as-if they are correct. This practice ignores model ambiguity, opens the door for misspecification of the decision problem, and leads to post-decision disappointment. We develop a framework to explore, evaluate, and optimize decision rules that explicitly account for the uncertainty in the first step estimation using statistical decision theory. We show how to operationalize our analysis by studying a stochastic dynamic investment model where the decision-makers take ambiguity about the model's transition dynamics directly into account.

JEL Codes D81, C44, D25

Keywords decision-making under uncertainty, robust Markov decision process

*Corresponding author: Philipp Eisenhauer, peisenha@uni-bonn.de. Philipp Eisenhauer is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1- 390838866, the TRA Modelling (University of Bonn) as part of the Excellence Strategy of the federal and state governments, and a postdoctoral fellowship by the AXA Research Fund. We thank Anton Bovier, Annica Gehlen, Lena Janys, Ken Judd, John Kennan, Simon Merkt, Sven Rady, Gregor Reich, John Rust, Jörg Stoye, and Rafael Suchy for numerous helpful discussions. 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. We gratefully acknowledge support by the AXA Research Fund.

Contents

1. Introduction

Decision-makers often confront uncertainties when determining their course of action. For example, individuals save to cover uncertain medical expenses in old age (French and Song, 2014). Firms set prices facing uncertainty about their competitive environment (Ilut, 2020), and policy-makers vote for climate change mitigation efforts facing uncertainty about future costs and benefits (Barnett et al., 2020). We consider the situation where a decision-maker posits a collection of economic models to inform his decision-making process. Each model formalizes the objectives and trade-offs involved. Within a given model, uncertainty is limited to risk as a model induces a unique probability distribution over possible futures. In addition, however, there is also ambiguity about the true model (Arrow, 1951; Knight, 1921).

In this context, we focus on the common practice in economics to estimate a subset of the model parameters outside the model and let the decision-makers inside the model treat these point estimates as-if they correspond to the true parameters.[1] This approach ignores ambiguity about the true model, resulting from the parametric uncertainty of the first-step estimation, and opens the door for the misspecification of the decision problem. As-if decision-makers, decision-makers who use the point estimates to inform decisions that would be optimal if the estimates were correct (Manski, 2021), face the risk of serious disappointment about their decisions. The performance of as-if decisions often turns out to be very sensitive to misspecification (Smith and Winkler, 2006), which is particularly consequential in dynamic models where the impact of erroneous decisions accumulates over time (Mannor et al., 2007). This danger creates the need for robust decision rules that perform well over a whole range of different models instead of an as-if decision rule that performs best for one particular model. However, increasing robustness, often measured by a performance guarantee under a worst-case scenario, decreases performance in all other scenarios. Striking a balance between the two objectives is challenging.

We develop a framework to evaluate as-if and robust decision rules in a decision-theoretic setting by merging insights from the literature on data-driven robust optimization (Bertsimas et al., 2018) and robust Markov decision processes (Ben-Tal et al., 2009) with statistical decision theory (Berger, 2010). We set up a stochastic dynamic investment model where the decision-maker takes ambiguity about the model's transition dynamics directly into account. We construct ambiguity sets for the transitions that are anchored in empirical estimates, statistically meaningful, and computationally tractable (Ben-Tal et al., 2013) using the Kullback-Leibler divergence

[1] See for example Berger and Vavra (2015), Blundell et al. (2016), Cagetti and De Nardi (2006), Chiappori et al. (2018), De Nardi et al. (2010), Ejrnæs and Jørgensen (2020), Fernández and Wong (2014), French and Jones (2011), Gourinchas and Parker (2002), Huo and Ríos-Rull (2020), Scholz et al. (2006), Sommer (2016), and Voena (2015).

(Kullback and Leibler, 1951). Our work brings together and extends research in economics and operations research to make econometrics useful for decision-making with models (Manski, 2021; Bertsimas and Thiele, 2006).

As an application, we revisit Rust (1987)'s seminal bus replacement problem which serves as a computational illustration in a variety of settings.[2] In the model, the manager Harold Zurcher implements a maintenance plan for a fleet of buses. He faces uncertainty about the future mileage utilization of the buses. To make his plan, he assumes that the mileage utilization follows an exogenous distribution and uses data on past utilization for its estimation. In the standard as-if analysis, the distribution is estimated in a first step and serves as a plug-in for the true unknown distribution. Harold Zurcher makes decisions as-if the estimate is correct and ignores any remaining ambiguity about future mileage utilization. We set up a robust version of the bus replacement problem to directly account for the estimation uncertainty and explore alternative decision rules' properties and relative performance.

In econometrics, there is burgeoning interest in assessing the sensitivity of findings to model or moment misspecification.[3] Our work is most closely related to Jørgensen (2021), who develops a measure to assess the sensitivity of results to fixing a subset of parameters of a model prior to the estimation of the remaining parameters. Our approach differs as we directly incorporate model ambiguity in the design of the decision-making process inside the model and assess the performance of a decision rule under misspecification of the decision environment. As such, our focus on ambiguity faced by decision-makers inside economic models draws inspiration from the research program summarized in Hansen and Sargent (2016) that tackles similar concerns with a theoretical focus. We complement recent work by Saghafian (2018), who works in a setting very similar to ours but does not use statistical decision theory to determine the optimal robust decision rule. In operations research, there exists a small number of applications of data-driven robust decision-making in a dynamic setting, which include portfolio allocation (Zymler et al., 2013), elective admission to hospitals (Meng et al., 2015), scheduling of liver transplantation (Kaufman et al., 2017), and the cost-effectiveness of colorectal cancer screening policies (Goh et al., 2018). To the best of our knowledge, none evaluates the performance of robust decisions against the as-if alternative in a decision-theoretic framework.

The structure of the remaining analysis is as follows. First, in Section 2, we present statistical decision theory as our framework to compare decision rules. We then set up a canonical model

[2]See for example Christensen and Connault (2019), Iskhakov et al. (2016), Reich (2018), and Su and Judd (2012).

[3]See for example Andrews et al. (2020), Andrews et al. (2017), Armstrong and Kolesár (2021), Bonhomme and Weidner (2020), Chernozhukov et al. (2020), Christensen and Connault (2019), and Honoré et al. (2020).

of a data-driven robust Markov decision process in Section 3 and outline the decision-theoretic evaluation of as-if and robust decision rules. Next, Section 4 presents our analysis of the robust bus replacement problem. The final section concludes.

2. Statistical decision theory

We now show how to compare as-if decision-making to its robust alternatives using statistical decision theory. We first review the basic setting and then turn to a classic urn example to illustrate some key points.

2.1. Decision problem

We study a decision problem in which the consequence $c \in \mathcal{C}$ of various alternative actions $a \in \mathcal{A}$ depend on the parameterization $\theta \in \Theta$ of an economic model. A consequence function $\rho : \mathcal{A} \times \Theta \mapsto \mathcal{C}$ details the consequence of action a under parameters θ:

$$c = \rho(a, \theta).$$

A decision-maker ranks consequences according to a utility function $u : \mathcal{C} \mapsto \mathbb{R}$ where higher values are more desirable. The structure of the decision problem $(\mathcal{A}, \theta, \mathcal{C}, \rho, u)$ is known, but there is uncertainty about the true parameterization θ_0, i.e., the consequences of a particular action remain ambiguous. However, an observed sample of data $\psi \in \Psi$ provides a signal about the true parameters as P_θ, the sampling distribution of ψ, differs by θ.

A statistical decision function $\delta : \Psi \mapsto \mathcal{A}$ is a procedure for determining an action for each possible realization of the sample. With as-if decision-making, the point estimates of the parameters $\hat{\theta}$ serve as a plug-in for the truth, ignoring their inherent uncertainty. This approach is just one particular example of a statistical decision function. We compare it against robust alternatives that explicitly account for the estimation uncertainty.

Statistical decision theory provides the framework to compare the performance of alternative decision rules $\delta \in \Gamma$. The utility achieved by δ is a random variable before the realization of ψ, and Wald (1950) suggests measuring performance for a given θ by its expected utility.

$$\mathrm{E}_\theta \left[u \left(\rho(\delta(\psi), \theta) \right) \right] = \int_\Psi u \left(\rho(\delta(\psi), \theta) \right) dP_\theta(\psi)$$

As the true parameterization is unknown, we need to aggregate the vector of expected utilities at each element in the parameter space.

In general, there is no single decision rule that yields the highest expected utility for all possible parameterizations and so determining the best decision rule δ^* is not straightforward. However, decision theory proposes various criteria (Gilboa, 2009; Marinacci, 2015). At the most fundamental level, any decision rule is admissible if another rule does not exist whose expected utility is always at least as high. In most cases, several decision functions are admissible, and additional optimality criteria are needed. We explore the common three: (1) maximin, (2) minimax regret, and (3) subjective Bayes.

The maximin decision (Gilboa and Schmeidler, 1989; Wald, 1950) is determined by computing the minimum performance for each decision rule at all points in the parameter space and choosing the one with the highest worst-case performance. Stated concisely,

$$\delta^* = \arg\max_{\delta \in \Gamma} \min_{\theta \in \Theta} \mathrm{E}_\theta \left[u(\rho\left(\delta(\psi), \theta\right)) \right].$$

For the minimax regret criterion (Manski, 2004; Niehans, 1948), we compute the maximum regret for each decision rule at all points in the parameter space and chooses the decision rule that minimizes the maximum regret. The regret of choosing a decision rule for any realization of θ is the difference between the maximum possible performance where the true parameterization informs the decision and its actual performance. Thus, the minimax regret criterion solves:

$$\delta^* = \arg\min_{\delta \in \Gamma} \max_{\theta \in \Theta} \underbrace{\left[\max_{a \in \mathcal{A}} u(\rho\left(a, \theta\right)) - \mathrm{E}_\theta \left[u(\rho\left(\delta(\psi), \theta\right)) \right] \right]}_{\text{regret}}.$$

Maximization of the performance under subjective Bayes (Savage, 1954) requires a subjective probability distribution f_θ over the parameter space. Then the alternative with the highest expected subjective utility is selected:

$$\delta^* = \arg\max_{\delta \in \Gamma} \int_\theta \mathrm{E}_\theta \left[u(\rho\left(\delta(\psi), \theta\right)) \right] df_\theta.$$

2.2. Urn example

We now illustrate the key ideas allowing us to compare as-if and robust decision-making using statistical decision theory with an urn example. As in our empirical application, we study a set of decision rules. We first compare the performance of two distinct rules and then determine the optimal rule within the set.

We consider an urn with black b and white w balls where the true proportion of black balls θ_0 is unknown. In this example, the action constitutes a guess $\tilde{\theta}$ about θ_0 after drawing a fixed number of n balls at random with replacement. The parameter and action space both correspond to the unit interval $\Theta = \mathcal{A} = [0, 1]$.

If the guess matches the true proportion, we receive a payment of one. On the other hand, in case of a discrepancy, the payment is reduced by the squared error. Thus, the consequence function takes the following form:

$$\rho(\tilde{\theta}, \theta_0) = 1 - (\tilde{\theta} - \theta_0)^2.$$

The sample space is $\Psi = \{b, w\}^n$ where a sequence $(b, w, b, \ldots, b)$ of length n is a typical realization of ψ. The observed number of black balls r among the n draws in a given sample ψ provides a signal about θ_0. The sampling distribution for the possible number of black balls R takes the form of a probability mass function (PMF):

$$\Pr(R = r) = \begin{pmatrix} n \\ r \end{pmatrix} (\theta_0)^r \, (1 - \theta_0)^{n-r}.$$

Any function $\delta : \{b, w\}^n \mapsto [0, 1]$ that maps the number of black draws to the unit interval is a possible statistical decision function.

We focus on the following class of decision rules $\delta \in \Gamma$:

$$\delta(r, \lambda) = \lambda \left(\frac{r}{n}\right) + (1 - \lambda) \left(\frac{1}{2}\right), \qquad \text{for some} \quad 0 \leq \lambda \leq 1.$$

The empirical proportion of black balls in the sample r/n provides the point estimate $\hat{\theta}$. The decision rule specifies the guess as a weighted average between the point estimate itself and the midpoint of the parameter space. The larger λ, the more weight is put on the point estimate. At the extremes, at $\lambda = 1$, the guess is the point estimate, while for $\lambda = 0$, the guess is fixed at 0.5.

We later determine the best decision rule within this class, but we start by comparing the performance of the two decision functions with $\lambda = 1$ and $\lambda = 0.9$. We refer to the former as the as-if decision rule as it announces the point estimate as-if it is the true one and $\lambda = 0.9$ as the robust decision rule for reasons that will become clear later. We evaluate their relative performance by aggregating the vector of expected payoffs over the unit interval using the different decision-theoretic criteria. We assume a linear utility function and thus directly refer to the monetary consequences of a guess as its payoff. Throughout, we set the number of draws n to 50.

Figure 1 shows the sampling distribution of R and the associated payoff for the two decision rules at $\theta = 0.4$. The robust decision rule outperforms the as-if rule for realizations of the point estimate smaller than its true value due to the shift towards 0.5. At the same time, the as-if rule leads to a higher payoff at the center of the distribution.

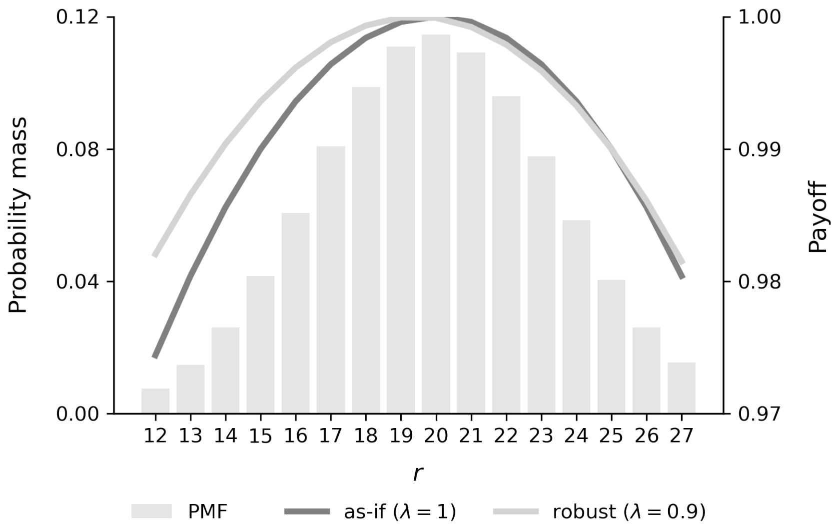
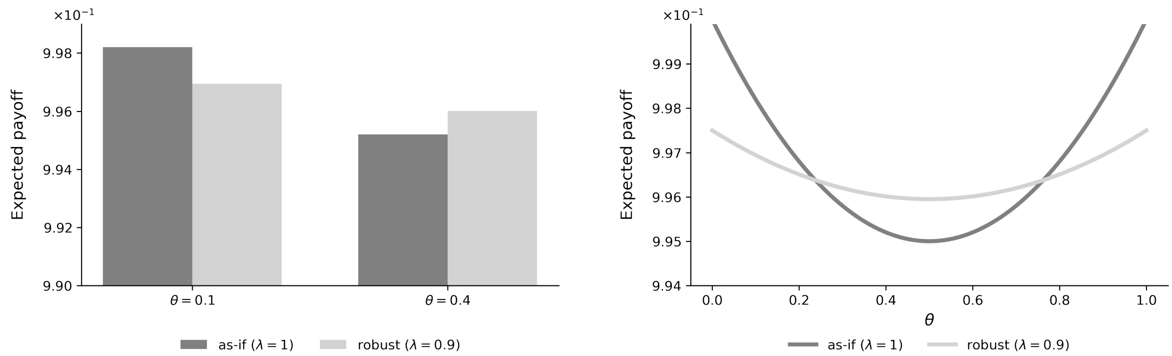


Figure 1.: Calculating the expected payoff

Figure 2 shows the expected payoff at different points in the parameter space. On the left, we show the expected payoff at two selected points. While the as-if decision rule performs better than the robust rule at $\theta = 0.1$, the opposite is true at $\theta = 0.4$. Thus, both decision rules are admissible as neither of the two outperforms the other for all possible true proportions. On the right, we trace out the expected payoff of both rules over the whole parameter space. The robust rule outperforms the as-if rule in the center of the parameter space but performs worse at the boundary. Overall, the performance of the robust rule is more balanced across the whole parameter space, which motivates its name.



(a) Expected payoff

(b) Performance

Figure 2.: Evaluating decision functions

Figure 3 ranks the two rules according to different decision-theoretic criteria. Both decision rules have their lowest expected payoff at $\theta = 0.5$. As the robust rule outperforms the as-if alternative at that point, the robust rule is preferred based on the maximin and minimax regret criteria. The maximin and minimax regret criteria are identical in this setting, as the payoff at the true proportion is constant across the parameter space. Using the subjective Bayes criterion with a uniform prior, however, we select the as-if decision rule as its better performance at the boundary of the parameter space is enough to offset its poor performance in the center.

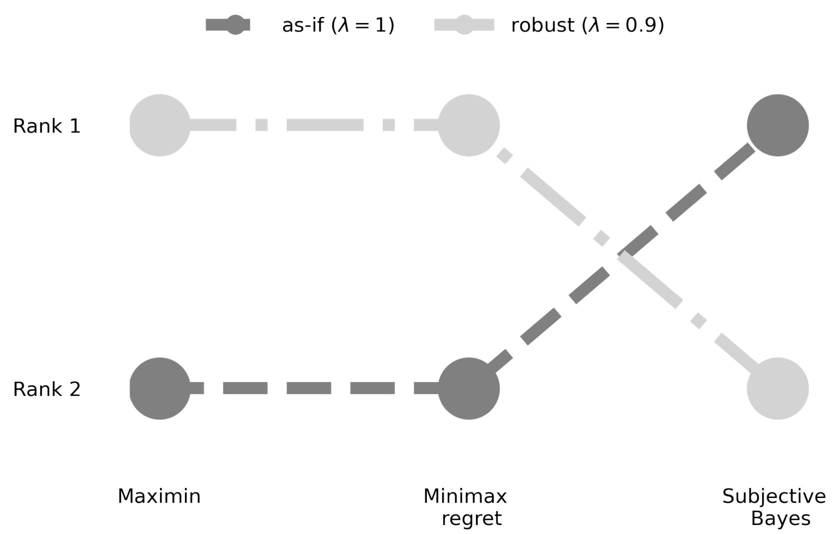
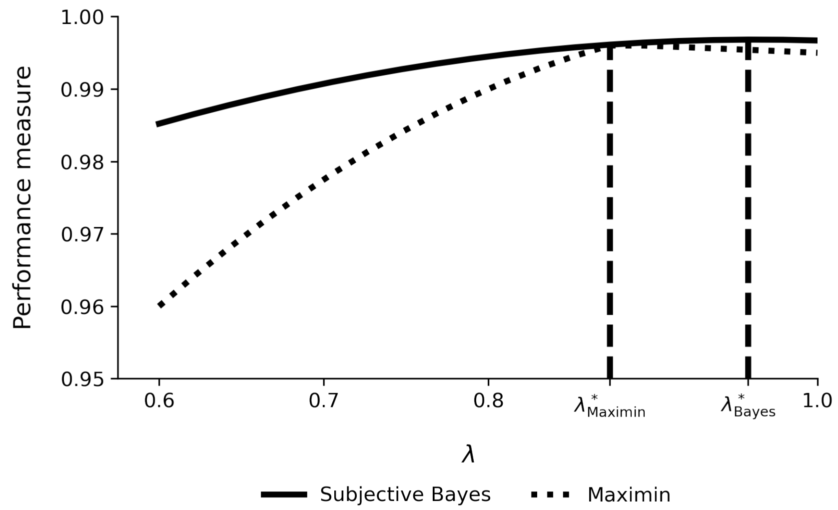


Figure 3.: Ranking of decision functions

Returning to the whole set of decision rules, we can construct the optimal rule for the alternative criteria by varying λ to maximize the relevant performance measure. For example, Figure 15 shows the optimal level of λ for the maximin and subjective Bayes criterion.



Notes: We omit the performance measure for the minimax regret criterion as $\lambda^*_{\text{Maximin}} = \lambda^*_{\text{Regret}}$ in this setting as noted earlier.

Figure 4.: Optimality of decision functions

The as-if decision rule ($\lambda = 1.0$) is not optimal under both criteria as at the optimum $0.8 <$

$\lambda^*_{\text{Maximin}} < \lambda^*_{\text{Bayes}} < 1$. The performance measure is more sensitive to the choice of λ under the maximin criterion than subjective Bayes.

In summary, the urn example illustrates the performance comparison of alternative decision rules over the whole parameter space of a model, and it shows how to construct optimal decision rules within a set of rules for alternative decision-theoretic criteria. Next, we move to the more involved setting of a sequential dynamic decision problem with ambiguous transitions and explore a class of statistical decision functions frequently used in the operations research literature.

3. Data-driven robust Markov decision process

We now outline our framework for the analysis of sequential decision-making in light of model ambiguity. We first present the general setup of a data-driven robust Markov decision process. We then turn to its solution approach, where we highlight the challenges introduced into the analysis when studying decision-making in this setting.[4] Finally, we outline the decision-theoretic evaluation of as-if decision-making and its robust alternatives.

Throughout, our exposition focuses on uncertainty in the transition dynamics of the Markov decision process. We do not address uncertainty about the parameters of the reward functions. In either case, our central insight to use statistical decision theory to explore, evaluate, and optimize robust decision rules applies. However, each setting introduces its unique computational challenges (Mannor and Xu, 2019). In line with our application, we discuss an infinite horizon model in discrete time, stationary utility and transition probabilities, and discrete states and actions.

3.1. Setting

At time $t = 0, 1, 2, \ldots$ a decision-maker observes the state of their environment $s_t \in \mathcal{S}$ and chooses an action a_t from the set of admissible actions $\mathcal{A}$. The decision has two consequences. It creates an immediate utility $u(s_t, a_t)$ and the environment evolves to a new state s_{t+1}. The transition from s_t to s_{t+1} is affected by the action, at least partly unknown, and governed by a transition probability distribution $p(s_t, a_t)$.

Decision-makers take the future consequences of the current action into account. A decision rule d_t specifies the planned action for all possible states within period t. A policy $\pi = \{d_0, d_1, d_2, \ldots\}$

[4] See Puterman (1994) for a textbook introduction to the standard MDP and Rust (1994) for a review of MDPs in economics and structural estimation.

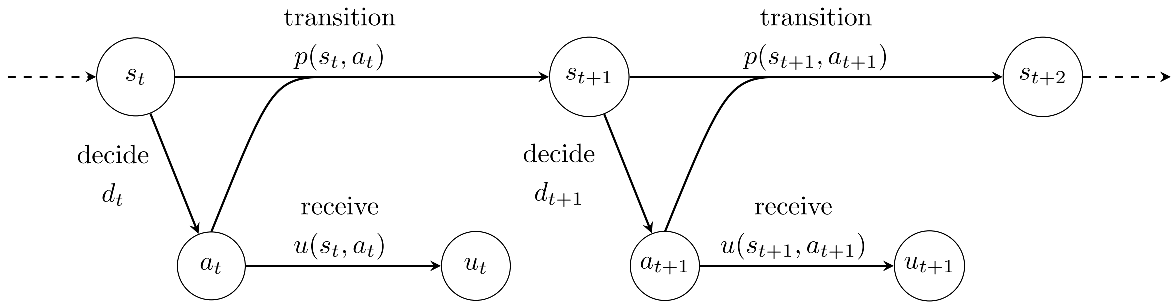


Figure 5.: Timing of events

is a collection of decision rules and specifies all actions for all time periods.

Figure 5 depicts the timing of events in the model. At the beginning of period t, a decision-maker learns about the utility of each alternative, chooses one of them according to the decision rule d_t, and receives its immediate utility. Then the state evolves from s_t to s_{t+1} and the process is repeated in $t+1$.

In a standard Markov decision process (MDP), a single transition probability distribution $p(s_t, a_t)$ is associated with each state-action pair. The decision-maker knows the distribution, and thus faces a decision problem under risk only. In a robust Markov decision process (RMDP) there is a whole set of distributions associated with each state-action pair collected in an ambiguity set $p(s_t, a_t) \in \mathcal{P}(s_t, a_t)$. The decision-maker only knows the ambiguity set and thus faces a risk for a given distribution and ambiguity about the true distribution. The MDP remains a special case of a RMDP when the ambiguity set is a singleton.

In a standard MDP, the objective of a decision-maker in state s_t at time t is to choose the optimal policy π^* from the set of all possible policies Π that maximizes their expected total discounted utility $\tilde{v}_t^{\pi^*}(s_t)$ as formalized in Equation (3.1):

$$\tilde{v}_t^{\pi^*}(s_t) \equiv \max_{\pi \in \Pi} \mathrm{E}^{\pi}\left[\sum_{r=0}^{\infty} \delta^{t+r} u(s_{t+r}, d_{t+r}(s_{t+r}))\right]. \tag{3.1}$$

The exponential discount factor δ parameterizes a taste for immediate over future utilities. The superscript of the expectation emphasizes that each policy induces a different probability distribution over sequences of possible futures. The standard value function $\tilde{v}_t^{\pi^*}(s_t)$ measures the performance of the optimal policy. This is true as long as transition probabilities used to construct the policy are in fact correct.

In an RMDP, the goal is to implement an optimal policy that maximizes the expected total discounted utility under a worst-case scenario. Given the ambiguity about the transition dynamics, a policy induces a whole set of probabilities over sequences of possible futures $\mathcal{F}^{\pi}$, and the worst-case realization determines its ranking. The formal representation of the decision-makers objective is Equation (3.2):

$$v_t^{\pi^*}(s_t) \equiv \max_{\pi \in \Pi} \left\{ \min_{\mathbf{P} \in \mathcal{F}^{\pi}} \mathrm{E}^{\mathbf{P}} \left[\sum_{r=0}^{\infty} \delta^{t+r} u(s_{t+r}, d_{t+r}(s_{t+r})) \right] \right\}. \tag{3.2}$$

We consider a setting where historical data provides information about the transition dynamics. In the data-driven standard MDP, the empirical probabilities $\hat{p}(s_t, a_t)$ serve as a plug-in for the truth. In a data-driven RMDP, they are used to construct the ambiguity sets for the transitions. We follow Ben-Tal et al. (2013) and create the ambiguity sets using statistical hypothesis testing. We restrict attention to distributions we cannot reject with a certain level of confidence $\omega \in [0, 1]$ around the empirical probabilities and collect them in an estimated ambiguity set $\hat{\mathcal{P}}(s_t, a_t; \omega)$.

Different values of ω result in different decision rules. Two special cases stand out. First, if $\omega = 0$, then a decision-maker treats the empirical probabilities as if they are correct. This case captures the notion of as-if decision making. Second, for $\omega = 1$, a robust decision-maker considers the worst-case scenario over the whole probability simplex at each state-action pair when constructing the optimal policy.

3.2. Solution

In the case of as-if decision-making, the goal is to maximize the expected total discounted utility as formalized in Equation (3.1). This requires evaluating the performance of all policies based on all possible sequences of utilities and the probability that each occurs. Fortunately, the stationary Markovian structure of the problem implies that the future looks the same whether the decision-maker is in state s at time t or any other point in time. The only variable which determines the value to the decision-maker is the current state s. Thus the optimal policy is stationary as well (Blackwell, 1965) and the same decision rule is used in every period. The value function is independent of time and the solution to the following Bellman equation:

$$\tilde{v}(s) = \max_{a \in \mathcal{A}} \left[u(s, a) + \delta \int \tilde{v}(s')\, \hat{p}(ds'|s, a) \right].$$

The as-if decision rule is recovered from $\tilde{v}(s)$ by finding the value $a \in \mathcal{A}$ that attains a maximum for each $s \in \mathcal{S}$.

Let $\mathbb{V}$ denote the set of all bounded real value functions on $\mathcal{S}$. Then the Bellman operator

$\tilde{\Gamma} : \mathbb{V} \to \mathbb{V}$ is defined as follows: For all $w \in \mathbb{V}$

$$\tilde{\Gamma}(w)(s) = \max_{a \in \mathcal{A}} \left[u(s,a) + \delta \int w(s')\, \hat{p}(ds'|s,a) \right], \quad s \in \mathcal{S}. \tag{3.3}$$

It allows to compute the value function $\tilde{v}(\cdot)$ as its unique fixed point (Denardo, 1967).

For a RMDP, where transition probabilities are ambiguous, the contraction mapping property of the Bellman operator and the optimality of a stationary deterministic Markovian decision rule requires the assumption of rectangularity of $\mathcal{F}^\pi$ (Iyengar, 2005; Nilim and El Ghaoui, 2005).

Rectangularity is a form of an independence assumption as the realization of any particular distribution in a state-action pair does not affect future realizations. The uncertainty is uncoupled across states and actions. This approach rules out any kind of learning about future ambiguity from past experiences due to, for example, a common source of uncertainty across states. While restrictive, most applications rely on the rectangularity assumption, as general notions of coupled uncertainties are intractable (Wiesemann et al., 2013).[5]

We now develop the formal definition of rectangularity. Let $\mathcal{M}(\mathcal{S})$ denote the set of all probability distributions on $\mathcal{S}$. Then the set of all conditional transition probability distributions associated with any decision rule d is given by:

$$\mathcal{F}^d = \{p : \mathcal{S} \to \mathcal{M}(\mathcal{S}) \,|\, \forall s \in \mathcal{S},\, p(s) \in \hat{\mathcal{P}}(s, d(s); \omega)\}.$$

For every state $s \in \mathcal{S}$, the next state can be determined by any $p \in \hat{\mathcal{P}}(s, d(s); \omega)$.

A policy π now induces a set of probability distributions $\mathcal{F}^\pi$ on the set of all possible histories $\mathcal{H}$. Any particular history $h = (s_0, a_0, s_1, a_1, \ldots)$ can be the result of many possible combinations of transition probabilities. Rectangularity imposes a structure on the combination possibilities.

Assumption 1. ***Rectangularity*** *The set $\mathcal{F}^\pi$ of probability distributions associated with a policy π is given by*

$$\begin{aligned}\mathcal{F}^\pi &= \left\{ \mathbf{P} \mid \forall\, h \in \mathcal{H} :\, \mathbf{P}(h) = \prod_{t=0}^{\infty} p(s_{t+1}|s_t, a_t), \text{ with } p(s_t, a_t) \in \hat{\mathcal{P}}(s_t, d_t(s_t); \omega) \text{ for } t = 0, 1, \ldots \right\} \\ &= \mathcal{F}^{d_0} \times \mathcal{F}^{d_1} \times \mathcal{F}^{d_2} \times \ldots = \prod_{t=0}^{\infty} F^{d_t},\end{aligned}$$

[5]See Mannor et al. (2016), and Goyal and Grand-Clement (2020) for recent attempts to introduce milder rectangularity conditions.

where the notation simply denotes that each element in $\mathcal{F}^{\pi}$ is a product of $p \in \mathcal{F}^{d_t}$, and vice versa (Iyengar, 2005).

Assumption 1 formalizes the idea that ambiguity about the transition probability distribution is uncoupled across states and time. All elements of the ambiguity sets can be freely combined to generate a particular history.

The objective when facing ambiguity is to implement a policy π^* that maximizes the expected total discounted utility under a worst-case scenario as presented in Equation (3.2). Under the rectangularity assumption, the decision-maker faces the same uncertainty, whether he is in state s at time t or any other point in time. Thus the value function is independent of time and solely depends on the current state s. It is the solution to the robust Bellman equation (3.4), where the future value is evaluated using the worst-case element in the ambiguity set (Iyengar, 2005).

$$v(s) = \max_{a \in \mathcal{A}} \left[u(s,a) + \delta \min_{p \in \hat{\mathcal{P}}(s,a;\omega)} \int v(s')\, p(ds'|s,a) \right]. \tag{3.4}$$

The robust Bellman operator on $\mathbb{V}$ follows directly: For all $w \in \mathbb{V}$

$$\Gamma(w)(s) = \max_{a \in \mathcal{A}} \left[u(s,a) + \delta \min_{p \in \hat{\mathcal{P}}(s,a;\omega)} \int w(s')\, p(ds'|s,a) \right] \quad s \in \mathcal{S}. \tag{3.5}$$

Algorithm 1 allows solving the RMDP by a robust version of the value iteration algorithm where κ denotes a convergence threshold. The calculation of future values under the worst-case scenario is the key difference to the standard approach.

Algorithm 1. Robust Value Iteration Algorithm

Input: $v \in \mathbb{V}, \kappa > 0$

For each $s \in \mathcal{S}$, set $\hat{v}(s) = \max_{a \in \mathcal{A}} \left\{ u(s,a) + \min_{p \in \hat{\mathcal{P}}(s,a;\omega)} \int v(s')\, p(ds'|s,a)] \right\}$.

while $\| v - \hat{v} \|_\infty > \kappa$ **do**

$v = \hat{v}$

$\forall\, s \in \mathcal{S}$, set $\hat{v}(s) = \max_{a \in \mathcal{A}} \left\{ u(s,a) + \min_{p \in \hat{\mathcal{P}}(s,a;\omega)} \int v(s')\, p(ds'|s,a)] \right\}$

end while

3.3. Evaluation

Robust policies are tailored to the simultaneous worst-case realization of all distributions in all ambiguity sets. The conservatism ensures a minimum performance over all distributions in the set. However, the performance of the robust rule in all other cases is disregarded. This indifference introduces a tradeoff when determining the size of the ambiguity set (Delage and Mannor, 2010). The larger the set, the more scenarios for which a minimum performance is ensured, but the decision rule's general performance suffers. This tradeoff is particularly pronounced when the actual structure of the decision problem exhibits coupled uncertainties that are ignored in the construction of the robust rule to ensure its computational tractability.

We use statistical decision theory to navigate the tradeoff. Each robust decision rule is a different statistical decision function, and we determine the optimal size of the ambiguity set ω^* using different decision-theoretic criteria. Adapting our urn example earlier to the setting of a data-driven robust Markov decision process, the parameter space corresponds to the set of transition probability distributions $\mathcal{L}(\mathcal{S}, \mathcal{A}) = \{p : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{M}(\mathcal{S})\}$. We observe data on the transition probabilities and measure the actual performance of a robust decision rule $\eta(\hat{p}; p_0, \omega)$ as the generated total discounted utility, which depends on the estimate of the transition probabilities $\hat{p}$, the true underlying probabilities p_0, and the confidence level ω. The standard decision-theoretic criteria translate to this setting as follows:

Maximin	$\omega^* = \arg\max_{\omega \in [0,1]} \min_{p \in \mathcal{L}(\mathcal{S},\mathcal{A})} \mathrm{E}_p \left[\eta(\hat{p}; p, \omega)\right]$
Minimax regret	$\omega^* = \arg\min_{\omega \in [0,1]} \max_{p \in \mathcal{L}(\mathcal{S},\mathcal{A})} \left[\max_{\tilde{\omega} \in [0,1]} \eta(p; p, \tilde{\omega}) - \mathrm{E}_p \left[\eta(\hat{p}; p, \omega)\right]\right]$
Subjective Bayes	$w^* = \arg\max_{\omega \in [0,1]} \int_{\mathcal{L}(\mathcal{S},\mathcal{A})} \mathrm{E}_p \left[\eta(\hat{p}; p, \omega)\right] \, df_p$

Note that even for genuinely uncoupled uncertainties, the maximin criterion does not automatically select the worst-case rule ($\omega = 1$). This particular rule is based on the worst-case scenario over the full probability simplex at each state-action pair. In fact, the worst-case rule might be inadmissible in particular settings as, for example, it is weakly dominated by the as-if rule. Suppose the true transitions correspond to the worst-case distributions. In that case, the distribution of sampled transitions is degenerate as the worst-case scenario at each state-action pair is the certain transition to the state with the lowest future value (Nilim and El Ghaoui, 2005). Thus, the as-if and worst-case rules share the same performance. For all other true transitions, the as-if rule may very well outperform the worst-case rule if the sampled data is sufficiently informative.

4. Robust bus replacement problem

We now set up a robust version of the seminal bus replacement problem and conduct our decision-theoretic analysis of alternative decision rules. First, we discuss the details of the computational implementation. Second, we conduct an ex-post analysis of as-if and robust decision rules based on Rust's (1987) original point estimates for the transition probabilities. Third, we implement an ex-ante decision-theoretic analysis. We evaluate the performance of as-if and robust decision rules over the whole probability simplex and determine the optimal size of the ambiguity set.

4.1. Setting

The bus replacement model is set up as a regenerative optimal stopping problem (Chow et al., 1971). We consider the sequential decision problem by a maintenance manager, Harold Zurcher, for a fleet of buses. He makes repeated decisions about their maintenance to maximize the expected total discounted utility under a worst-case scenario. Each month t, a bus arrives at the bus depot in state $s_t = (x_t, \epsilon_t)$ described by its mileage since the last engine replacement x_t and other signs of wear and tear ϵ_t. He is faced with the decision to either conduct a complete engine replacement ($a_t = 1$) or perform basic maintenance work ($a_t = 0$). The cost of maintenance $c(x_t)$ increases with the mileage state, while the cost of replacement RC remains constant. In the case of an engine replacement, the mileage state is reset to zero.

The immediate utility of each action is given by:

$$u(a_t, x_t) + \epsilon_t(a_t) \quad \text{with} \quad u(a_t, x_t) = \begin{cases} -RC & a_t = 1 \\ -c(x_t) & a_t = 0. \end{cases}$$

Harold Zurcher makes his decisions in light of uncertainty about next month's state variables captured by their conditional distribution $p(x_t, \epsilon_t, a_t)$.

Although in this framework the utility and consequently the value function is finite in each state, they are not uniformly bounded. This property, however, is a crucial assumption for the results of Blackwell (1965) and Denardo (1967) on the contraction property of the Bellman operator and the stationarity of the optimal policy. For the original as-if-analysis, Rust (1988) circumvents this problem by imposing conditional independence between the observable and unobservable state variables, i.e. $p(x_{t+1}, \epsilon_{t+1}|x_t, \epsilon_t, a_t) = p(x_{t+1}|x_t, a_t)\, q(\epsilon_{t+1}|x_{t+1})$, and assuming that the unobservables $\epsilon_t(a_t)$ are independent and identically distributed according to an extreme value distribution with mean zero and scale parameter one. These two assumptions, together with the additive separability between the observed and unobserved state variables in the immediate utilities, ensure that the expectation of the next period's value function is independent of the time. The regenerative structure of the process implies, that the transition probabilities in case of replacement in any mileage state correspond to the probabilities of maintenance in mileage state 0. Therefore, the expected value function is the unique fixed point of a contraction mapping on the reduced space of mileage states only. In addition, the conditional choice probabilities $P(a_t|x_t)$ have a closed-form solution (McFadden, 1973). We build on these results and extend them to our robust setting with ambiguous transition dynamics. The proof is available in Appendix A.

In the analysis of the original bus replacement problem, the distribution of the monthly mileage transitions are estimated in a first step and used as plug-in components for the subsequent analysis. We extend the original setup and explicitly account for the ambiguity in the estimation. Following the arguments on the regenerative structure of the process above, we can incorporate ambiguity in the decision-making process with ambiguity sets conditional on the mileage states x only. We construct ambiguity sets $\hat{\mathcal{P}}(x;\omega)$ based on the Kullback-Leibler divergence D_{KL} (Kullback and Leibler, 1951) that are anchored in empirical estimates $\hat{p}(x)$, statistically meaningful, and computationally tractable.

Our ambiguity set takes the following form for each mileage state x:

$$\hat{\mathcal{P}}(x;\omega) = \left\{ q \in \mathring{\Delta}_{|J_x|} : D_{KL}(q \parallel \hat{p}(x)) = \sum_{i=1}^{|J_x|} q_i \ln\left(\frac{q_i}{\hat{p}(j_i|x)}\right) \leq \rho_x(\omega) \right\},$$

where $J_x = \{j_1, \ldots, j_{|J_x|}\}$ denotes the set of all states, which have an estimated non-zero probability to be reached from x, $\mathring{\Delta}_{|J_x|} = \{p \in \mathbb{R}^{|J_x|} \,|\, p_i > 0 \text{ for all } i = 1, \ldots, |J_x| \text{ and } \sum_{i=1}^{|J_x|} p_i = 1\}$ is the interior of the $(|J_x| - 1)$ - dimensional probability simplex, and $\rho_x(\omega)$ captures the size of the set for the state x with a given level of confidence ω.

Iyengar (2002) and Ben-Tal et al. (2013) provide the statistical foundation to calibrate $\rho_x(\omega)$ such that the true (but unknown) distribution p_0 is contained within the ambiguity set for a given level of confidence ω. Let χ^2_{df} denote a chi-squared random variable with df degrees of freedom and let $F_{df}(\cdot)$ denote its cumulative distribution function with inverse $F^{-1}_{df}(\cdot)$. Then, the following approximate relationship holds as the number of observations N_x for state x tends to infinity (Pardo, 2005):

$$\begin{aligned} \omega &= \Pr[\, p_0 \in \hat{\mathcal{P}}(x;\omega)\,] \\ &\approx \Pr[\, \chi^2_{|J_x|-1} \leq 2N_x\rho_x(\omega)\,] \\ &= F_{|J_x|-1}(2N_x\rho_x(\omega)). \end{aligned}$$

Therefore we can calibrate the size of the ambiguity set based on the following relationship:

$$\rho_x(\omega) = \tfrac{1}{2N_x} F^{-1}_{|J_x|-1}(\omega). \tag{4.1}$$

We use Rust (1987)'s original data to inform our computational experiments. His data consists of monthly odometer readings x_t and engine replacement decisions a_t for 162 buses. The fleet consists of eight groups that differ in their manufacturer and model. We focus on the fourth group of 37 buses with a total of 4,292 monthly observations. We discretize mileage into 78 equally spaced bins of length $5,000$ and set the discount factor to $\delta = 0.9999$.

Figure 6 highlights the limited information about the true distribution of mileage utilization. It shows the number of observations available to estimate next month's utilization for different levels of accumulated mileage. While there are more than 1,150 observations on buses with less than 50,000 miles, there are only about 220 with more than 300,000.

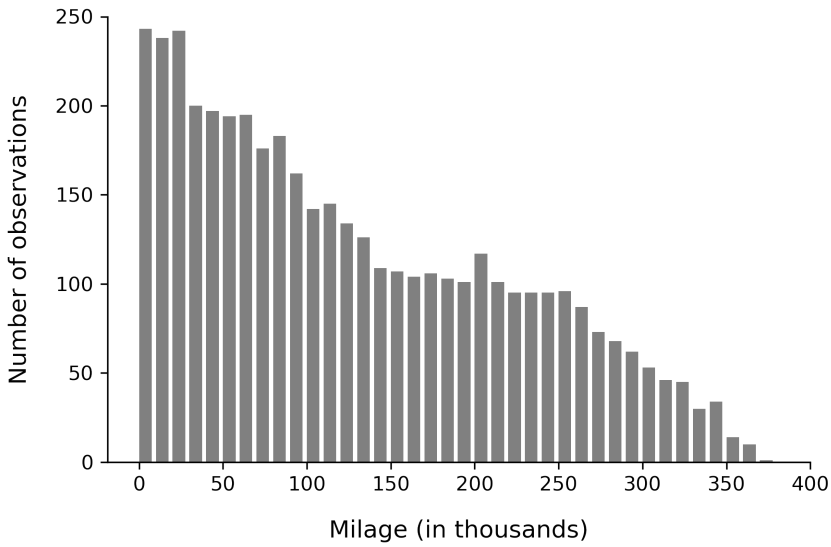


Figure 6.: Distributions of observations

We analyze a specific example of Rust (1987)'s bus replacement problem. We do not use his reported estimates of the maintenance and replacement costs. Given these estimates, Harold Zurcher's decisions are mainly driven by the unobserved state variable ϵ_t, and so ambiguity about the evolution of the observed state variable x_t does not have a substantial effect on decisions. We ensure that a bus's accumulated mileage has a considerable impact on the timing of engine replacements by increasing the maintenance and replacement costs compared to their empirical estimates. Thus, we specify the following cost function $c(x_t) = 0.4\, x_t$ and set the replacement costs RC to 50.

We solve the model using a modified version of the original nested fixed point algorithm (NFXP) (Rust, 1988). We determine the worst-case transition probabilities in each successive approximation of the fixed point. Given the size of the ambiguity set, we can determine the worst-case probabilities as the solution to a one-dimensional convex optimization problem (Iyengar, 2005; Nilim and El Ghaoui, 2005).[6]

4.2. Ex-post analysis

We now present the estimated transition probabilities and the worst-case distributions. We construct as-if and robust policies with varying confidence levels, outline the resulting differences in maintenance decisions, and evaluate their relative performance under different scenarios.

Figure 7 shows the point estimates $\hat{p}$ for the transition probabilities of monthly mileage usage. We pool all 4,292 observations to estimate this distribution by maximum likelihood, and thus the probability of next's period mileage utilization is the same for each state x_t. We only observe increases of at most $J = 3$ grid points per month. For about 60% of the sample, monthly bus utilization is between 5,000 to 10,000 miles. Very high usage of more than 10,000 miles amounts to only 1.2%.

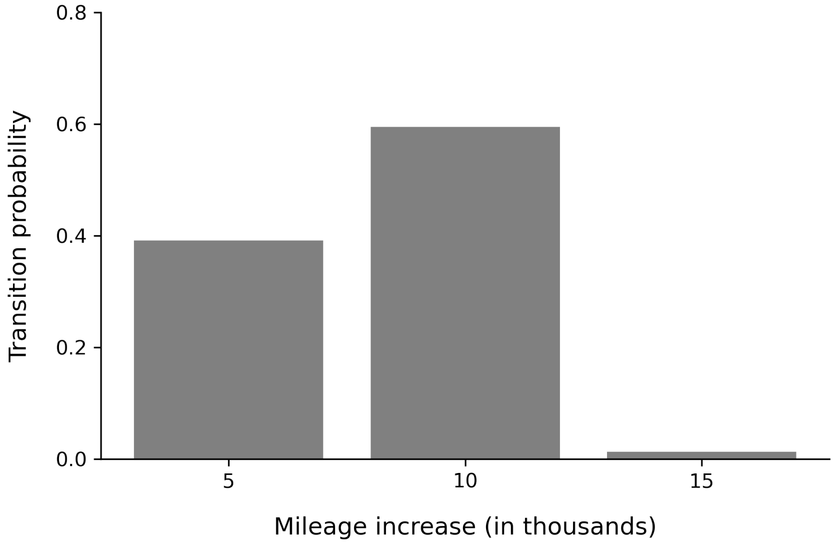


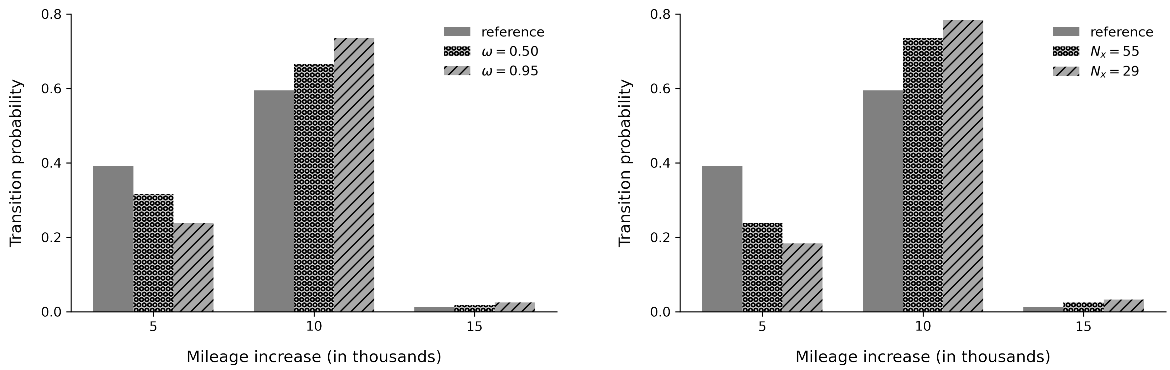
Figure 7.: Estimated transition probabilities

The confidence level ω and the available number of observations N_x determine the size of the ambiguity set as outlined in Equation (4.1). From now on, we mimic state-specific ambiguity sets by constructing them based on the average number of 55 observations per state. Note that

[6]The core routines are implemented in our group's ruspy (2020) and robupy (2020) software packages and publicly available.

while the estimated distribution is the same for all mileage levels, its worst-case realization is not. However, there are only minor differences across mileage levels, so we focus our following discussion on a bus with an odometer reading of 75,000.

Figure 8 shows the transition probabilities for different sizes of the ambiguity set. We vary the confidence level for the whole number of observations ($N_x = 55$) on the left, while on the right, the level of confidence remains fixed ($\omega = 0.95$) and we cut the number of observations roughly in half. The larger the ambiguity set, the more probability is attached to higher mileage utilization, resulting in higher costs overall. For example, while the probability of mileage increases of 10,000 or more is an infrequent occurrence in the data, its probability increases first to 1.7%. It then doubles to 2.5% as we increase the confidence level. When only about half the data is available, this probability increases even further to 3.2%.



(a) Variation in ω, ($N_x = 55$) **(b)** Variation in N_x, ($\omega = 0.95$)

Figure 8.: Worst-case transition probabilities

Harold Zurcher chooses whether to perform regular maintenance work on a bus or replace its complete engine each month. The assumed transition probabilities correspond to their worst-case transitions within the ambiguity set. So any differences between the as-if and worst-case distributions translate into different maintenance policies.

Figure 9 shows the maintenance probabilities for different levels of accumulated mileage and alternative policies. Overall, the maintenance probability decreases with accumulated mileage as maintenance gets more costly than an engine replacement. Robust policies result in a higher probability of maintenance compared to the as-if policy. Under the worst-case transitions, a bus is more likely to experience higher usage during the period. As maintenance cost is determined by the mileage level at the beginning of the period, maintenance becomes more attractive. For example, again considering a bus with 75,000 miles, the as-if maintenance probability is 25% while it is 33% ($\omega = 0.50$) and 43% ($\omega = 0.95$) following the robust policies.

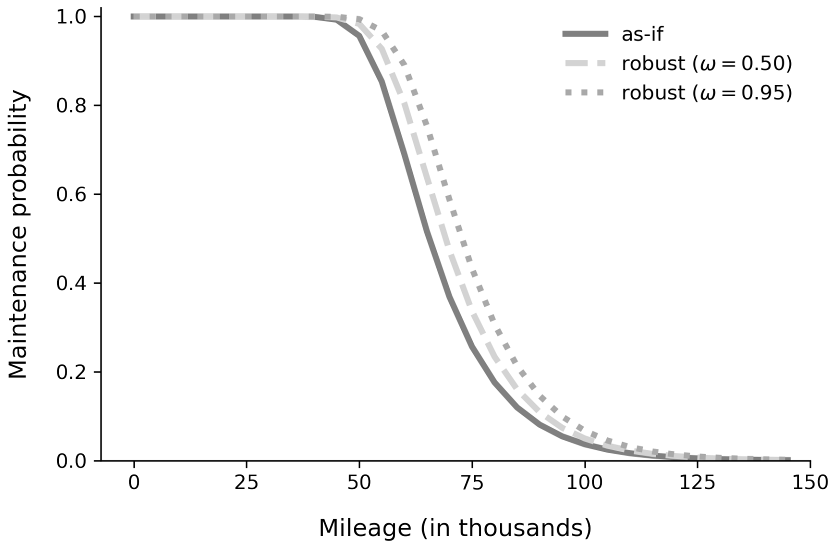


Figure 9.: Maintenance probabilities

To gain further insights into the differences between the as-if and robust policies, we simulate a fleet of 1,000 buses for 100,000 months under the alternative policies.

Figure 10 shows the level of accumulated mileage over time for a single bus under different policies. It clarifies our simulation setup, where we apply other policies to the same bus. The realizations of observed transitions and unobserved signs of wear and tear remain the same. The bus accumulates more and more mileage until Harold Zurcher replaces the complete engine, and the odometer resets to zero. The first replacement happens after 20 months at 60,000 miles following the as-if policy, while it is delayed for another four months under the robust alternative ($\omega = 0.95$). As its timing differs, the odometer readings will start to diverge after 20 months, even though monthly utilization remains the same.

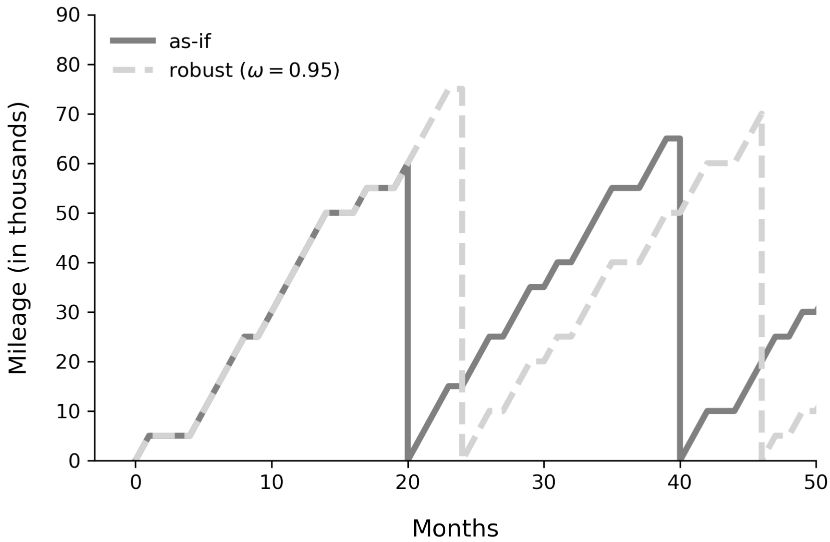


Figure 10.: Single bus under alternative policies

We now evaluate the as-if and robust policies at the boundary of the ambiguity set. We measure the performance of the alternative decision rules based on their total discounted utility under different assumed and actual mileage transitions.

Figure 11 shows the performance of the as-if policy over time when the worst-case distribution for a confidence level of 0.95 governs the actual transitions. It illustrates the sensitivity of the as-if policy to perturbations in the transition probabilities. The solid line corresponds to its expected long-run performance without misspecification of the decision problem, while the dashed line indicates its observed performance. After about 20,000 months, it accumulates the expected long-run average cost and performs about 14% worse overall.

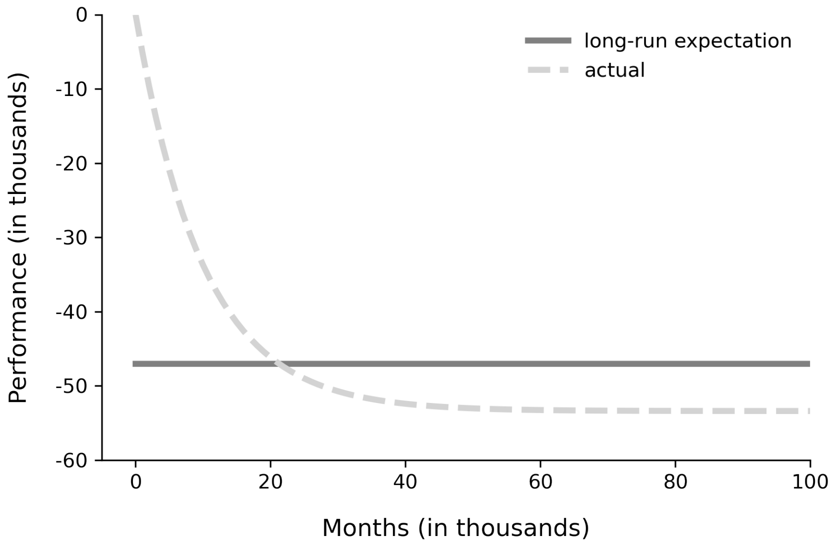
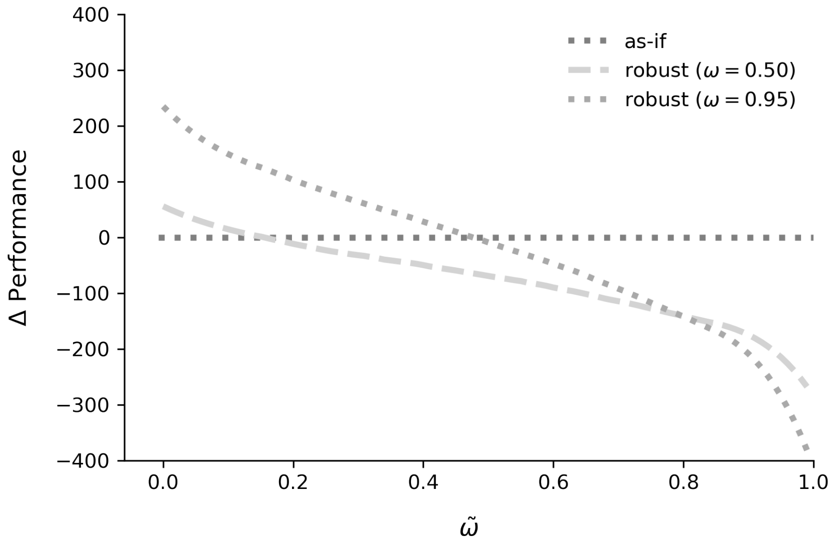


Figure 11.: Performance of as-if policy

Figure 12 shows the average difference in performance between the as-if and two robust policies with a confidence level of 0.50 and 0.95. The actual transitions follow the worst-case distribution with varying ω. A positive value indicates that the robust policy outperforms the as-if policy. In the absence of any misspecification, the as-if policy must defeat any other policy. The same is true for the robust policies when the actual transitions are governed by the same ω used for their construction. Nevertheless, the as-if policy continues to outperform both robust policies for moderate levels of ω. For worst-case distributions with ω larger than 0.2, the first robust policy ($\omega = 0.5$) starts to beat the as-if policy. For the other robust policy ($\omega = 0.95$), this is true for worst-case transitions of ω equal to 0.5.



Notes: We apply a Savitzky-Golay filter (Savitzky and Golay, 1964) to smooth the simulation results.

Figure 12.: Performance and misspecification

4.3. Ex-ante analysis

We now turn to a decision-theoretic evaluation of the ex-ante performance of as-if and robust decision-making over the whole probability simplex.

We operationalize our analysis as follows. In line with Rust's (1987) assumption on the distribution of the mileage utilization, we specify a uniform grid with 0.1 increments over the interior of the two-dimensional probability simplex $\mathring{\Delta}_3$. At each grid point, we draw 100 samples of 55 random mileage utilizations. For each sample, we construct decision rules for a grid $\omega = \{0.0, 0.1, \ldots, 1.0\}$ using the estimated transition probabilities. The uncertainties are coupled across states in our setting as the same underlying probability creates the sample of bus utilizations. The rectangularity assumption does not reflect the economic environment, but is imposed for the construction of the robust decision rules to ensure tractability. We then simulate the rules' actual performance and compute their expected performance by averaging across the 100 runs for each grid point. Using this information, we measure the different rules' performance based on the maximin criterion, the minimax regret rule, and the subjective Bayes approach using a uniform prior.

The computational burden is considerable even for our relatively simple application. At each grid point in the probability simplex, we solve 100 robust Markov decision processes for each robust decision rule. However, the analysis is amenable to parallelization using modern high-performance computational resources (Dongarra and Van der Steen, 2012) as we can process each estimated transition probability independently. We use standard linear interpolation between the grid points.

In Figure 13 we illustrate the differences in expected performance between a robust decision rule ($\omega = 0.1$) against the as-if alternative over the probability simplex.

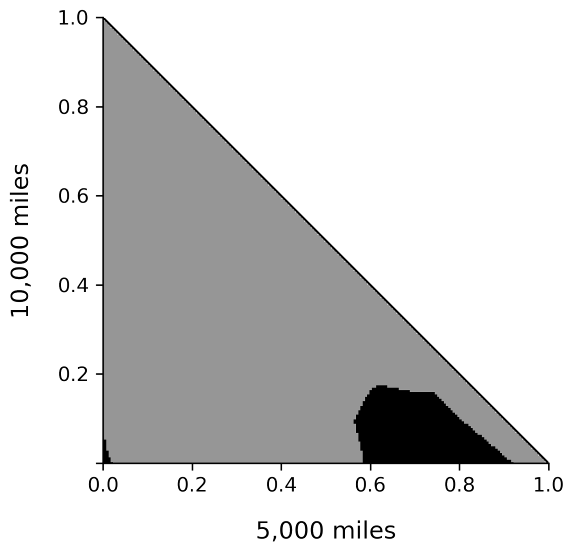


Figure 13.: Relative performance of decision functions

In the gray areas, the as-if decision rule outperforms the robust alternative based on its expected performance. The opposite is true for the black areas. A robust decision rule performs very well when the true probability of mileage increases of $5,000$ per month is high and of $10,000$ low. Otherwise, the as-if decision rule outperforms the robust alternative. Thus no rule dominates the other and it is essential to aggregate the performance over the whole probability simplex using decision theory before settling on a decision rule.

Figure 14 ranks the as-if decision rule against selected robust alternatives for the different performance criteria.

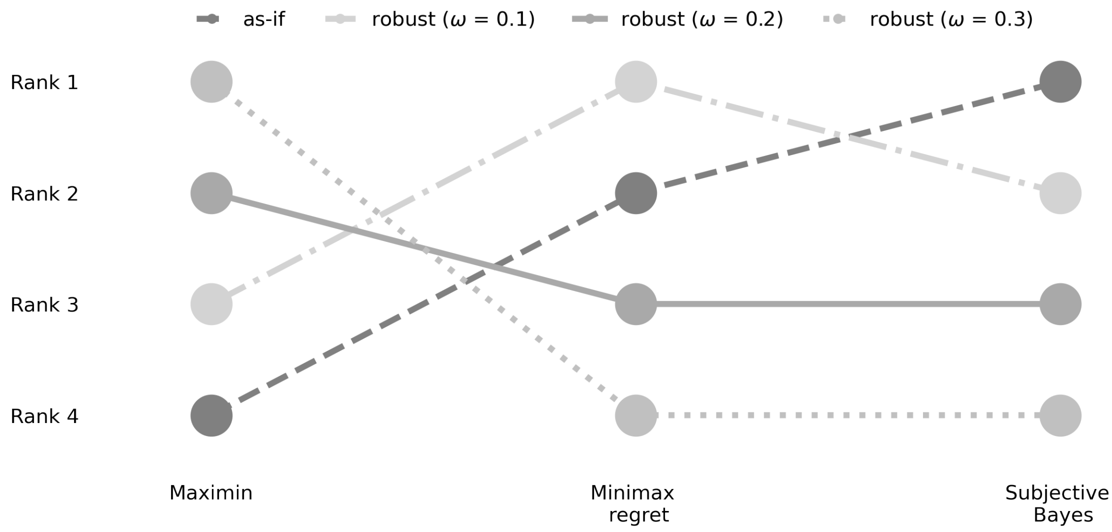


Figure 14.: Ranking of decision functions

Based on a maximin criterion, the decision rules rank the higher, the greater the confidence level ω. The robust rule with $\omega = 0.3$ comes in first, while as-if decisions rank last. Thus, decision-makers can improve their worst-case outcomes by adopting a robust decision rule. However, this comes at a cost, as indicated by the improved rankings for the as-if decision rule as we move to different criteria. As-if decisions move to second place for minimax regret. When we aggregate performance across all states using a subjective Bayes approach, the as-if rule comes first based on a subjective Bayes assessment. Thus, our approach clarifies the trade-offs involved when choosing a particular decision function.

We now determine the optimal size of the ambiguity set ω^* for each decision-theoretic criteria. Figure 15 shows the minimum performance of the robust decision rules for varying levels of ω normalized between zero and one. Among all rules, the robust rule with $\omega = 0.36$ has the highest minimum performance. It thus strikes a balance between the conservatism of the worst-case approach and the protection against unfavorable transition probabilities. Based on the maximin criterion, the as-if decision rule performs worst.

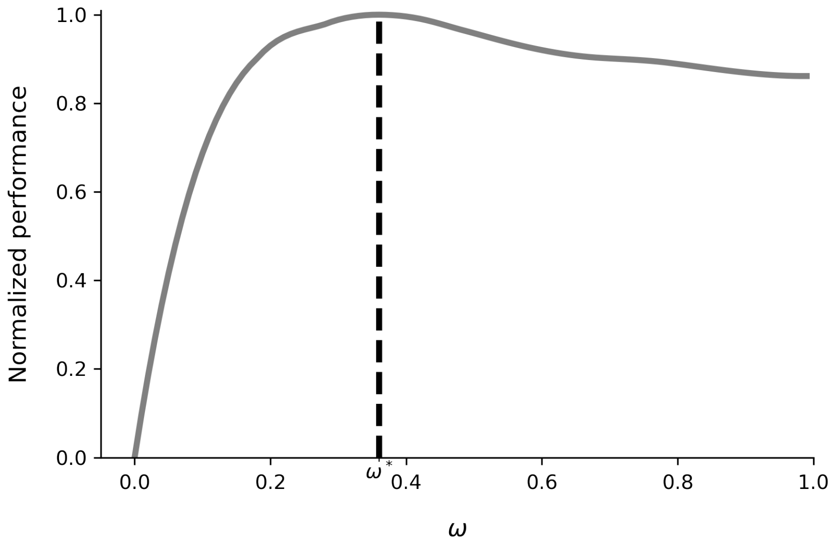


Figure 15.: Optimality of decision functions

The minimax regret criterion leads to a slightly reduced level of $\omega = 0.1$. The as-if decision rule is optimal based on the subjective Bayes criterion.

5. Conclusion

Economists often estimate a subset of their model parameters outside the model and let the decision-makers inside the model treat these point estimates as-if they are correct. This approach ignores model ambiguity. We set up a stochastic dynamic investment model where the decision-maker faces ambiguity about the model's transition dynamics. We propose a framework to evaluate decision rules that ignore the ambiguity against alternatives that take it directly into account. We show how to determine the optimal level.

As our core contribution, we combine ideas from data-driven robust optimization (Bertsimas et al., 2018), robust Markov decision processes (Ben-Tal et al., 2009) and statistical decision theory (Berger, 2010) to enable decision-making with models under uncertainty (Manski, 2021). This insight transfers directly to many other settings. For example, the COVID-19 pandemic provides a timely example of economists informing policy-making using highly parameterized models in light of ubiquitous uncertainties (Avery et al., 2020). When analyzing these models, economists treat numerous of their parameters as-if they are known. However, their actual values are uncertain as they are often estimated based on external data sources. Our research illustrates how to conduct robust policy-making and evaluate its relative performance against policies that ignore uncertainty using statistical decision theory. Such an approach promotes a sound decision-making process as it provides decision-makers with the tools to navigate and communicate the uncertainties they face in a systematic fashion (Berger et al., 2021).

References

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.

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

Avery, C., Bossert, W., Clark, A., Ellison, G., and Ellison, S. F. (2020). Policy implications of models of the spread of coronavirus: Perspectives and opportunities for economists. *NBER Working Paper*.

Barnett, M., Brock, W., and Hansen, L. P. (2020). Pricing uncertainty induced by climate change. *The Review of Financial Studies*, 33(3):1024–1066.

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.

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

Berger, D. and Vavra, J. (2015). Consumption dynamics during recessions. *Econometrica*, 83(1):101–154.

Berger, J. O. (2010). *Statistical decision theory and Bayesian analysis*. Springer, New York City, NY.

Berger, L., Berger, N., Bosetti, V., Gilboa, I., Hansen, L. P., Jarvis, C., Marinacci, M., and Smith, R. D. (2021). 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.

Bertsimas, D. and Thiele, A. (2006). Robust and data-driven optimization: Modern decision making under uncertainty. In Johnson, M. P., Norman, B., and Secomandi, N., editors, *Models, Methods, and Applications for Innovative Decision Making*, pages 95–122. INFORMS.

Blackwell, D. (1965). Discounted dynamic programming. *The Annals of Mathematical Statistics*, 36(1):226–235.

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

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

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, frictions, and wealth. *Journal of Political Economy*, 114(5):835–870.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., and Robins, J. M. (2020). Locally robust semiparametric estimation. *arXiv Working Paper*.

Chiappori, P.-A., Dias, M. C., and Meghir, C. (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy*, 126(S1):26–72.

Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*. Houghton Mifflin, Boston, MA.

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

De Nardi, M., French, E., and Jones, J. B. (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, 118(1):39–75.

Delage, E. and Mannor, S. (2010). Percentile optimization for Markov decision processes with parameter uncertainty. *Operations Research*, 58(1):203–213.

Denardo, E. V. (1967). Contraction mappings in the theory underlying dynamic programming. *SIAM Review*, 9(2):165–177.

Dongarra, J. J. and Van der Steen, A. J. (2012). High-performance computing systems: Status and outlook. *Acta Numerica*, 21:379–474.

Ejrnæs, M. and Jørgensen, T. H. (2020). Family planning in a life-cycle model with income risk. *Journal of Applied Econometrics*, 35(5):567–586.

Fernández, R. and Wong, J. C. (2014). Divorce risk, wages and working wives: A quantitative life-cycle analysis of female labour force participation. *The Economic Journal*, 124(576):319–358.

French, E. and Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732.

French, E. and Song, J. (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, 6(2):291–337.

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

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

Goh, J., Bayati, M., Zenios, S. A., Singh, S., and Moore, D. (2018). Data uncertainty in Markov chains: Application to cost-effectiveness analyses of medical innovations. *Operations Research*, 66(3):697–715.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Goyal, V. and Grand-Clement, J. (2020). Robust Markov decision process: Beyond rectangularity. *arXiv Working Paper.*

Hansen, L. P. and Sargent, T. J. (2016). *Robustness.* Princeton University Press, Princeton, NJ.

Honoré, B., Jørgensen, T., and de Paula, Á. (2020). The informativeness of estimation moments. *Journal of Applied Econometrics*, 35(7):797–813.

Huo, Z. and Ríos-Rull, J.-V. (2020). Sticky wage models and labor supply constraints. *American Economic Journal: Macroeconomics*, 12(3):284–318.

Ilut, C. (2020). Paralyzed by fear: Rigid and discrete pricing under demand uncertainty. *Econometrica*, 88(5):1899–1938.

Iskhakov, F., Lee, J., Rust, J., Schjerning, B., and Seo, K. (2016). Comment on "Constrained optimization approaches to estimation of structural models". *Econometrica*, 84(1):365–370.

Iyengar, G. N. (2002). Robust dynamic programming. *CORC Tech Report.*

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

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

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

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

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

Mannor, S., Mebel, O., and Xu, H. (2016). Robust MDPs with k-rectangular uncertainty. *Mathematics of Operations Research*, 41(4):1484–1509.

Mannor, S., Simester, D., Sun, P., and Tsitsiklis, J. N. (2007). Bias and variance approximation in value function estimates. *Management Science*, 53(2):308–322.

Mannor, S. and Xu, H. (2019). Data-driven methods for Markov decision problems with parameter uncertainty. In Netessine, S., editor, *Operations Research & Management Science in the Age of Analytics*, pages 101–129. INFORMS.

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

Manski, C. F. (2021). Econometrics for decision making: Building foundations sketched by Haavelmo and Wald. *Econometrica*, forthcoming.

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

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York City, NY.

Meng, F., Qi, J., Zhang, M., Ang, J., Chu, S., and Sim, M. (2015). A robust optimization model for managing elective admission in a public hospital. *Operations Research*, 63(6):1452–1467.

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

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

Pardo, L. (2005). *Statistical inference based on divergence measures.* Chapman & Hall, London, UK.

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

Reich, G. (2018). Divide and conquer: Recursive likelihood function integration for hidden Markov models with continuous latent variables. *Operations Research*, 66(6):1457–1470.

robupy (2020). A Python package for robust optimization.

ruspy (2020). An open-source package for the simulation and estimation of a prototypical infinite-horizon dynamic discrete choice model based on Rust (1987).

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

Rust, J. (1988). Maximum likelihood estimation of discrete control processes. *SIAM Journal on Control and Optimization*, 26(5):1006–1024.

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.

Saghafian, S. (2018). Ambiguous partially observable Markov decision processes: Structural results and applications. *Journal of Economic Theory*, 178:1–35.

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

Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639.

Scholz, J. K., Seshadri, A., and Khitatrakun, S. (2006). Are americans saving "optimally" for retirement? *Journal of Political Economy*, 114(4):607–643.

Smith, J. E. and Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322.

Sommer, K. (2016). Fertility choice in a life cycle model with idiosyncratic uninsurable earnings risk. *Journal of Monetary Economics*, 83:27–38.

Su, C.-L. and Judd, K. L. (2012). Constrained optimization approaches to estimation of structural models. *Econometrica*, 80(5):2213–2230.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review*, 105(8):2295–2332.

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

Wiesemann, W., Kuhn, D., and Rustem, B. (2013). Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183.

Zymler, S., Kuhn, D., and Rustem, B. (2013). Worst-case value at risk of nonlinear portfolios. *Management Science*, 59(1):172–188.

A. The robust contraction mapping

Rust (1987) shows that the expectation of the next period's value function is a fixed point on the mileage states x only. He uses the regenerative property of the mileage process and introduces a separate notion $\tilde{EV}(x)$ for the expected value function of maintenance. $\tilde{EV}(x)$ is the fixed point of the contraction mapping defined as follows: For all $v \in \mathbb{V}$

$$\tilde{\Gamma}(v)(x) = \sum_{x' \in X} p(x'|x) \log \sum_{a \in \{0,1\}} \exp\big(u(x', a) + \delta\, v((1-a)\, x')\big),\; x \in X. \tag{A.1}$$

We adopt a similar approach, following Iyengar (2005), and show:

Theorem 1. *Let the robust Bellman operator $\Gamma : \mathbb{V} \to \mathbb{V}$ be defined as follows: For all $v \in \mathbb{V}$*

$$\begin{aligned}\Gamma(v)(x) &= \int \max_{a \in \{0,1\}} \Big[u(x,a) + \epsilon(a) + \delta \min_{p \in \mathcal{P}((1-a)\,x,\,\omega)} \sum_{x' \in X} p(x')v(x')\Big] q(d\epsilon) \\ &= \log \sum_{a \in \{0,1\}} \exp \Big[u(x,a) + \delta \min_{p \in \mathcal{P}((1-a)\,x,\,\omega)} \sum_{x' \in X} p(x')v(x')\Big]. \end{aligned} \tag{A.2}$$

Then $\Gamma(\cdot)$ is a contraction mapping on $\left(\mathbb{V}, \|\cdot\|_\infty\right)$ with unique fixed point EV.

Proof. Let $v, w \in \mathbb{V}$ be arbitrary. Fix $x \in X$ and assume without loss of generality that $\Gamma(w)(x) \geq \Gamma(v)(x)$. Let $\nu > 0$ be arbitrary. Then choose $p^a \in \mathcal{P}((1-a)x, \omega)$, such that

$$\begin{aligned}&\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \min_{p \in \mathcal{P}((1-a)\,x,\,\omega)} \sum_{x' \in X} p(x')v(x')] \geq \\ &\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p^a(x')v(x')] - \nu.\end{aligned}$$

By construction:

$$\begin{aligned}&\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \min_{p \in \mathcal{P}((1-a)\,x,\,\omega)} \sum_{x' \in X} p(x')w(x')] \leq \\ &\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p^a(x')w(x')].\end{aligned}$$

Rust (1988) shows for any conditional distribution measure p and mileage state $x \in X$:

$$\begin{aligned}&\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p(x')w(x')] - \max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p(x')v(x')] \\ &\leq \delta \max_{a \in \{0,1\}} |\sum_{x' \in X} p(x')(w(x') - v(x')| \leq \delta \, \|w - v\|_\infty .\end{aligned}$$

This holds in particular for p^a, which yields:

$$\begin{aligned}
0 &\leq \Gamma(w)(x) - \Gamma(v)(x) \\
&\leq \int \Bigg(\max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p^a(x') w(x')] \\
&\qquad - \max_{a \in \{0,1\}} [u(x,a) + \epsilon(a) + \delta \sum_{x' \in X} p^a(x') v(x')] + \nu \Bigg) q(d\epsilon) \\
&\leq \int \left(\delta \, \|w - v\|_\infty + \nu \right) q(\epsilon) \\
&= \delta \, \|w - v\|_\infty + \nu.
\end{aligned}$$

Arguing vice versa for $\Gamma(w)(x) \leq \Gamma(v)(x)$, this implies that

$$\|\Gamma(w) - \Gamma(v)\|_\infty \leq \delta \, \|w - v\|_\infty + \nu.$$

With ν arbitrary and $\delta \in [0,\, 1)$ this shows that Γ is a contraction mapping on $\mathbb{V}$ with respect to $\|\cdot\|_\infty$. As $\left(\mathbb{V},\, \|\cdot\|_\infty\right)$ is a Banach space, the result is established. □