Abstract of "Modeling and Simulation of Artificial Societies to Study Precarity and Inequity" by Pegah Nokhiz, Ph.D., Brown University, May 2024.

Financial insecurity characterizes modern life. It reflects ongoing instability in employment and earnings which are notably pronounced in the gig economy. This instability is compounded by the widespread use of automated decision-making tools that directly affect employment and income. Over time, this "precariousness" unfolds as a sequence of events for individuals. Thus, to understand and address it, a shift in perspective from decision-makers to individuals is necessary. This requires that we develop "artificial societies" – computational simulations of an agent-based behavioral model capable of capturing various related phenomena simultaneously, including individual consumption responses to financial shocks, the influence of predictive tools on income, and the long-term behavior of individuals striving to maximize utility. This individual-level perspective is one direction to study precarity and inequity in artificial societies with computational simulations or models designed to replicate and investigate the intricate behavior of complex social systems. However, there is also a societal-level viewpoint wherein neither a singular decision-maker nor defined agent behavior rules exist. Consequently, there is a need for a model that does not attempt to describe underlying systems or capture individual actions. Adopting a system-based approach to studying inequity in feedback loops opens avenues to explore social systems that are otherwise challenging to model directly. This dissertation first introduces the concept of latent financial instability, or precarity, to the artificial intelligence community. It develops agent-based computational models embodying realistic human-like behaviors to explore precarity dynamics, drawing from various strands of inquiry in economics. Additionally, we investigate work schedule instability and the impact of foresight on financial security. Next, we present a model from linear systems theory to quantify feedback in social systems holistically, enabling the examination of long-term policy effects even without individually characterized feedback mechanisms. Our frameworks facilitate the examination of precarity dynamics, the development of mitigation strategies for precarity, policy investigations, and the production and sustainment of long-term equity.

Modeling and Simulation of Artificial Societies to Study Precarity and Inequity

by
Pegah Nokhiz
B. S., Shahid Beheshti University, 2015
Sc. M., University of Kansas, 2018

A dissertation submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
in the Department of Computer Science at Brown University

Providence, Rhode Island
May 2024

This dissertation by Pegah Nokhiz is accepted in its present form by
the Department of Computer Science as satisfying the dissertation requirement
for the degree of Doctor of Philosophy.

Date ______________ ________________________________
Suresh Venkatasubramanian, Advisor

Recommended to the Graduate Council

Date ______________ ________________________________
Stephen Bach, Reader

Date ______________ ________________________________
Neal Patwari, Reader
Washington University in St. Louis

Approved by the Graduate Council

Date ______________ ________________________________
Thomas A. Lewis
Dean of the Graduate School

Vita

Pegah Nokhiz completed her Bachelor of Science in Computer Science at Shahid Beheshti University and her Master of Science in Computer Science at the University of Kansas. She started her Ph.D. at the University of Utah in 2018 and transferred to Brown University in 2021. At Brown, she continued her research with her PhD advisor Professor Suresh Venkatasubramanian as an affiliate of Brown's Center for Technological Responsibility, Reimagination and Redesign (CNTR) at the Data Science Institute. At Brown, her research work has focused on the long-term effects of automated decision-making (on artificial societies), fairness in machine learning, bridging quantitative and qualitative methods from other disciplines by analyzing and quantifying interdisciplinary notions from moral philosophy and social sciences (to gauge the ethical and social impacts of AI tools), and econometric approaches to simulating individual behavior.

Acknowledgements

First, I extend my gratitude to my advisor, Professor Suresh Venkatasubramanian, for his steadfast guidance and support during every step of my PhD journey. I have been very fortunate to be his student. His mentorship, patience, and willingness to explore unconventional research avenues, allowing me the freedom to delve into my research interests, have been invaluable to me. I am also grateful for his role in shaping my research interests and providing me with countless opportunities for growth.

I am deeply thankful for the invaluable contributions, support, and constructive feedback provided by my collaborators throughout the entirety of my work. Their involvement has been indispensable, and I feel privileged to have had such exceptional collaborators. In particular, I would like to extend my gratitude to Professors Vivek Srikumar and Aditya Bhaskara for their guidance and mentorship. I am also grateful to Aravinda Kanchana Ruwanpathirana for his patience and substantial input into our joint endeavors. Furthermore, I extend my thanks to Vivek Gupta for his backing and motivation in our exploration of various research avenues.

I also thank my committee members, Professors Neal Patwari and Stephen Bach, for their feedback and support. Additionally, I am grateful to my colleagues at the University of Utah and Brown University for creating a supportive and welcoming environment.

Last but foremost, I wish to express my profound appreciation to my parents for their steadfast love and support throughout this journey. Their unwavering encouragement has been indispensable in realizing my aspirations for a PhD.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Financial insecurity in the U.S. is on the rise, accelerated by the growth of the gig economy and the associated income instability, increasing inequality, and the effects of algorithmic decision-making [79, 207]. Studying *financial insecurity* necessitates a shift in focus – from the point of view of the decision-maker to that of the decision subject. This centering of the subject on individuals is a direction that unlocks the importance of moving away from aggregate measures to examine the long-term effects of decision-making on people [148].

Conversely, closed systems with a specific decision-maker and certain engagement rules in place do not suffice to examine broad-level societal dynamics. That is, if the objective is to achieve equity in a broader societal system, studying the system in isolation is insufficient. In a societal-level system, a singular decision-maker or exact agent behavior rules are non-existent. Additionally, analysis of societal systems can be complicated by the presence of *feedback*, in which historical and current inequities influence future inequity [170].

Our goal in this work is to study these concepts in an artificial society as computer simulations or computational models designed to imitate and investigate the behavior of complex social systems [22]. These simulated societies replicate the intricate dynamics of individuals to gain insights into how individuals behave. Further, by offering a way to examine the effects of regulations and other interventions, they shed light on the complex dynamics of society as a whole.

To account for both perspectives, in this dissertation, we explore two world views in studying financial instability and inequity in societies. This primarily consists of four previous conference publications and manuscripts [148, 170, 149, 150]. First, we can study instability from the individuals' perspective when subjected to stimuli and financial shocks. Individuals react by making decisions on how to spend and save their money – an "agent-based" or "individual-level" perspective (Chapter 3 [148], Chapter 4 [149], and Chapter 5 [150]). Secondly, we can examine the high-level complex economy and dynamics of power and not go deeper into modeling individual behavior (or modeling feedback mechanisms individually and in their full complexity). This decidedly broad-level perspective is "societal-level" (Chapter 6 [170]).

1.1.1 Individual (Agent-based) Perspective: Financial Instability[1]

Financial insecurity is a characteristic of modern life in America [207]. A combination of socioeconomic factors, increasing inequality, unstable jobs, and data-driven algorithmic decision-making has left families increasingly vulnerable to financial "shocks" that can have an outsized and often irreversible long-term effect on their finances.

Latent socioeconomic factors. Some individuals and households are more vulnerable financially due to latent socioeconomic factors. According to financial reports [79, 207], these latent factors can be demographic – communities of color tend to be more likely to experience negative income shocks – as well as economic, with factors ranging from the need to support family, an unstable job situation, and so on.

Job instability. The rise of the gig economy has led to an increase in paycheck instability with associated long-term effects. Consider a gig worker versus an office worker with a stable job and salary. While both might start with the same set of observable economic features, i.e., similar assets and income levels, the gig worker's finances (income and employment status) are prone to be more volatile in the long run due to the nature of short-term and unpredictably valued contracts in the gig economy [43].

The compounding effect of (negative) decisions. In an economy where decisions about individual finances are increasingly controlled by algorithms, the effects of adverse decisions can compound over time, yielding disproportionately large variances in outcomes even if individuals

[1]Adapted from: I. Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana, Neal Patwari, and Suresh Venkatasubramanian. "Precarity: Modeling the Long Term Effects of Compounded Decisions on Individual Instability." In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, pp. 199-208. 2021.
II. Pegah Nokhiz*, Aravinda Kanchana Ruwanpathirana*, Neal Patwari, and Suresh Venkatasubramanian. "Agent-based Simulation of Decision-making under Uncertainty to Study Financial Precarity", In Proceedings of PAKDD 2024.

have very similar financial starting points. The role of algorithmic decision-making in hiring and assessing office workers has been well-documented [167, 21, 121]. In addition, gig work is another battleground where the effect of algorithmic decisions become significant. For example, customer satisfaction and job acceptance ratings play a pivotal role on gig platforms like Uber and Lyft as they are used to incentivize workers to perform to companies' accepted standards [114, 168]. One or two low ratings and marginal rating differences get picked up by the algorithm that sets assignments and rates for drivers. Over time, these differences can get compounded and bring down a worker's chances of being employed again for a new gig by the algorithm. This is particularly true for new workers without a long history of high ratings [114, 175].

The looming risk of financial ruin. A recent CNBC survey revealed that 58% of American households are living paycheck to paycheck [40]. A household that cannot build up savings is one that is vulnerable to (even small) financial shocks. At the margins, this precariousness can tip households over the edge into bankruptcy and even homelessness: a recent study from UCSF indicated that the most commonly reported cause of homelessness was a loss of or reduction in income (with 12% of respondents indicating this) [108].

The study of the social impact of automated decision-making has focused largely on issues of fairness at the point of decision, evaluating the fairness (with respect to a population) of a sequence or pipeline of decisions, or examining the dynamics of a game between the decision-maker and the decision subject [219]. Ripple effects are not captured by changes in income or wealth alone or by one decision alone. To study financial insecurity, we must reorient our frame of reference away from the decision-maker and towards the decision subject; away from aggregates of decisions over a population and towards aggregates of decisions (for an individual) over time.

There is a name for this *precarious* state of being and the missing piece is an examination of precarity: a transdisciplinary term coined by Judith Butler [19, 25] which characterizes the latent instability, *precariousness*, and therefore vulnerability of people's lives. Researchers have proposed quantitative measures of precarity [173]. Precarity also depicts how negative decisions can have ripple effects on one's well-being [19, 25] and it has been linked to automated decision-making [148].

Overview. In this dissertation we model and study this missing piece. That is, we propose modeling frameworks to simulate the effects of automated decisions on an individual over time, incorporating a quantification of their precarity. In an individual-level (agent-based) view of financial instability (captured by precarity), we are interested in exploring the following research questions:

- How do income shocks affect the precarity of individuals over time? Does algorithmic decision-making exacerbate the precarity of individuals?

- Does precarity affect various income classes non-uniformly?
- Can we study interventions' effects which would be difficult if not impossible to test "in the wild"? Do the timing and amount (span) of interventions matter?
- Do latent factors other than observable features contribute to long-term precarity?
- How can we improve existing agent-based consumption models to study precarity?

The framework we propose to study and answer the questions above draws on a number of interrelated threads of work in economics that seek to model human consumption. We first introduce precarity to the AI community and develop rational utility maximization as well as bounded rationality computational models to capture individual-level financial behavior under uncertainty at each step of the simulation. We then improve our framework by incorporating human-analogous features like minimum consumption requirements, time of death, and the desire to avoid bankruptcy (in Chapters 3 and 4).

Contributions. Our contributions on an individual-level view are listed as follows:

- We first introduce the idea of precarity to the artificial intelligence community and automated decision-making systems.
- By quantifying precarity, we show how the underlying population and various income classes get more precarious in a long-term compounded decision setting using a newly designed simulation framework.
- We improve existing agent-based econometric models (and thus our simulation framework) on long-term financial behavior by adding realistic constraints such as risk of bankruptcy, consumption constraints, and time of death to them.
- Our results illustrate how precarity, if ignored by policy-makers, can exacerbate the ill-effects of automated decision-making, i.e., precarity can be exacerbated by the compounding effects of repeated algorithmic decisions that take financial variables into account when making predictions that in turn cause future financial shocks.
- We study various mitigation strategies to mitigate precarity.

1.1.2 Individual (Agent-based) Perspective: Temporal Instability[2]

In addition to low wages and limited benefits, economically disadvantaged workers often contend with unpredictable work schedules and there is a tendency to overlook this temporal aspect. Previous research reveals that 80% of food industry workers have minimal to no influence on their schedules, and 69% are obligated by the system to maintain schedules that are "open and available" for work at any given time [125, 188]. 8.4% of workers and consultants/contractors aged 18-65 reported significant fluctuations in their income on a monthly basis. 51% of these workers with volatile work timetables attributed their income instability to an irregular work schedule [134].

With a just-in-time work schedule, managing one's finances becomes increasingly difficult. Workers earn income through their employment and use it to fulfill daily necessities like food, shelter, transportation, clothes, recreation, and so on. In this context, the income earned by workers is allocated to different purposes, leading to corresponding gains in utility based on how much they save and consume optimally. Thus, planning a life (and maintaining the corresponding financial welfare) with a volatile schedule would be an exceedingly difficult task.

Overview. Given that a) work schedule instability has a direct impact on individuals' employment and income, and b) the affected groups often comprise part-time workers, individuals in lower-asset/income/education categories, and women of color who receive unfair advance notice compared to their counterparts [134, 188], important lines of inquiry are:

- What are the adverse effects of this scheduling discrepancy on these protected groups in precarious work environments?
- How can these potential adverse impacts be effectively measured?
- Recently, there have been new policy regulations and measures proposed that mandate employers offer more advance notice when establishing or altering work schedules, aiming to enhance predictability for workers [70]. How do these interventions work empirically?

To answer these questions (in Chapter 5), we need to study the consumption and saving behavior of workers in the face of unforeseen financial shocks from unstable work timetables. Thus, the primary goal of this chapter in the dissertation is to thoroughly investigate the dynamics and effects of work instability on the earned utility of workers. This goal necessitates the development of an agent-based behavioral model capable of simultaneously simulating several interconnected phenomena, i.e., the model should encompass how individual consumption responds to unexpected financial shocks and

[2]Adapted from: Pegah Nokhiz*, Aravinda Kanchana Ruwanpathirana*, Aditya Bhaskara, and Suresh Venkatasubramanian. "Counting Hours, Counting Losses: The Toll of Unpredictable Work Schedules on Financial Security", currently under submission.

the behavior of individuals as they seek to maximize utility with different levels of information on future events.

Most economic models of consumption and savings assume that individuals possess the ability to fully look into the future when making consumption decisions [37, 35]. On the contrary, in real life, information is limited, and our ability to act on it is also limited. Thus, how would individuals' financial decision-making change if they only had a limited window of information into how their financial well-being might change? This is particularly relevant in settings where workers are increasingly only given limited visibility into income/asset-affecting decisions like how many hours they will work or at what rate they will earn.

Therefore, we seek to model this using the language of online learning, i.e., an adaptive update of the workers' consumption policies (an online learning paradigm) where the policy is recalculated at each step as more information on work schedules becomes available. We investigate how far utility maximizing strategies depend on the degree of "foresight".

Contributions. Overall, the main contributions of this temporal aspect are:

- We propose a novel algorithm, capable of handling varying levels of lookahead.
- We carry out a formal and empirical analysis to show that workers who possess a lookahead benefit from an advantage that increases proportionally with the magnitude of their lookahead.
- We explore temporal equity, particularly in the context of the implications of lack of advance notice (future lookahead) on work timetables that affect the disadvantaged subpopulations more acutely.
- We explore various intervention strategies (adopted from fair workplace laws and acts) to examine the adverse effects of just-in-time work schedules.

Note on our agent-based work. Our methodology in the "agent-based" or "individual-level" perspective (Chapter 3 [148], Chapter 4 [149], and Chapter 5 [150]) employs simulations as a toolkit to study long-term behavior. The efficacy of a simulation depends on the quality of the models utilized to build it, and we employ rational agents for our epistemic inquiries. However, the use of rational agents does not imply that we assume individuals always act optimally and are entirely rational in reality; rather, it acknowledges that even if they are, there are limits on what they can do. That is, by inquiry into limits, we show even under ideal models of utility maximization, the lack of predictability and lookahead has concrete consequences in terms of precarity and financial stability.

1.1.3 Societal Perspective: Inequity and Feedback[3]

Why have societal inequities endured despite decades of activism, educational efforts, policy reforms, and the professed values of equality and non-discrimination? For example, segregation and inequities in housing and employment have persisted in the U.S. despite decades under the Fair Housing Act and Equal Employment Opportunity laws [29]. The answer, in our view, is rooted in the phenomenon of *feedback*. In every system in which inequity persists over time, there are feedback mechanisms which enable it to survive – as *1984* posits, "The object of power is power" [157]. Conversely, activism, public pressure, and equitable policies are used to push toward equity – Frederick Douglass said "If there is no struggle, there is no progress" [44] – and these can be seen as reactions to historical and present inequality, and thus are also a type of feedback.

This dissertation argues that feedback modeling tools from systems theory are helpful in quantitatively modeling mechanisms of feedback that help perpetuate and combat inequity. Good models help us gain more understanding of the processes maintaining the status quo and can inform policies which "produce and sustain equity" [105] when deployed in the real world. However, the economy and dynamics of power are complex, and we do not intend to model the feedback mechanisms individually and in their full complexity (as is attempted in system dynamics [88]). Instead, we focus on inequity at a systems level, essentially from the outside of a black box, both maintained and diminished over time by feedback mechanisms which quantify how much it will change or stay stationary. What is the benefit of such a model? For one, we can use system identification tools to find quantitative estimates for each type of feedback, and compare the amount of feedback by type in different systems. Further, we can use the model to estimate future inequity. Finally, new policies and algorithms which influence future inequity can be modeled as having feedback mechanisms which operate in parallel with those in society and their impact on equity estimated.

Feedback is not merely a systemic response to how a system evolves over time. In societal settings, policies and interventions form another form of feedback. When new algorithms operate in parallel with societal mechanisms, as depicted in Figure 1.1, we can model their combined impact and forecast how they impact our trajectory towards equity.

Overview. We consider the following questions:

- How can we design a framework for feedback that captures a wide variety of societal feedback mechanisms?

[3] Adapted from: Lydia Reader, Pegah Nokhiz, Cathleen Power, Neal Patwari, Suresh Venkatasubramanian, and Sorelle Friedler. "Models for understanding and quantifying feedback in societal systems." In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, pp. 1765-1775. 2022.

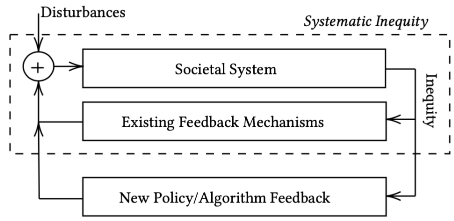


Figure 1.1: Current societal systems have feedback mechanisms which make group inequities persist over time. New policies & algorithms can have feedback mechanisms that act in parallel, altering the inequity over time.

- How well does our model forecast future inequity? And how does the model compare to existing simple models?
- How can we use the model to understand the effects of new policies or algorithms?

We employ one of the most extensively studied feedback systems in systems theory which is the PID framework [113], consisting of proportional, integral and derivative forms of feedback. In this dissertation, we use the PID framework purely as a descriptive tool (and not as a controller as is common in systems theory) to answer the questions above. We argue with extensive examples and a formal analysis that this simple framework for feedback captures a wide variety of societal feedback mechanisms [170] (in Chapter 6).

Contributions. We can summarize the contributions of this societal perspective as follows.

- We present a method for modeling feedback in societal systems based on the PID framework from (linear) systems theory.
- We show with an extensive list of examples the ways in which the PID framework effectively captures real-world examples of moves toward (and away from) equity.
- We demonstrate the working of this model using three case studies involving historical and persistent inequity.
- We demonstrate how the model can be used to evaluate the effects of policy shifts and interventions.

1.2 Dissertation Organization

The remainder of the dissertation is structured as follows.

- Chapter 2 on background and related work introduces the main econometric, social, agent-based modeling/simulation, and feedback modeling concepts used in other chapters.
- Chapter 3 introduces a new framework to study precarity. We first introduce the idea of precarity and model the long-term effects of compounded decisions on individual instability using consumption models.
- Chapter 4 expands the previous chapter on agent-based modeling to study precarity with realistic constraints. In this chapter, we develop a realistic model with realistic constraints to investigate latent factors contributing to precarity and mitigation strategies.
- Chapter 5 explores work schedule instability in precarious work environments from a temporal perspective. We explain the motivation behind employing an online model to study individual-level behavior on how acquiring future information (e.g., the timetable for the upcoming workweeks) would help with gaining more utility in the long run (both formally and empirically).
- In Chapter 6, we transition to a societal perspective, departing from the individual-level focus of previous chapters. We explore feedback in societal-level systems wherein our proposed framework is employed for understanding and quantifying feedback in societal systems using PID.
- Lastly, in Chapter 7 we summarize the overall contributions and explain possible future directions.

Chapter 2

Background and Related Work

We draw on ideas from economics, algorithmic fairness, social sciences, and the literature on simulations of social systems. In this chapter, we will provide an overview of the key background information pertinent to our contributions in the subsequent chapters (adapted from [148, 170, 149, 150]). We will commence by delineating the concept of precarity and methods for its quantification. Subsequently, we will delve into the broader context of algorithmic fairness and simulation. Next, we explain models of consumption and savings, the relevant work on temporal instability in work schedules, and the literature on system dynamics.

2.1 Precarity

Precarity [18, 25] is a multi-faceted concept that very broadly speaks to the instability and *precariousness* of people's lives. It has been interpreted as an economic condition [213, 34], a sociological condition that speaks the interconnectedness and therefore vulnerability of human lives [25, 26], as a descriptor of a political class characterized by irregular or transient employment [195, 67] or as a psychological condition of exclusion and displacement [5]. We can also interpret precarity as the instability associated with sequences of negative decisions: specifically, the way in which repeated negative outcomes can increase the likelihood of one falling into poverty. Ritschard et al. [173] were the first to attempt to quantify precarity (in the context of the labor market) by looking at transitions between more or less precarious states (for example, a full time versus a part time job). This work observes that negative transitions have the most critical role in increasing precarity. Aneja et al. [6] study the effect of incarceration on access to credit – arguing that incarceration reduces the access to credit, which in turn increases recidivism. Another relevant work is by Abebe et al. [3]. In it, they build a theoretical model to capture the effect of *income shocks* on one's chance of going bankrupt and propose efficient allocations of limited stimulus to maximize the expected number of individuals saved from bankruptcy.

2.1.1 Quantifying Precarity

Precarity (as discussed above) is a broad interdisciplinary notion describing the instability of modern life. In this chapter, we focus on the *economic* aspects of precarity – how financial and other shocks create uncertainty around one's financial status. Within the social sciences, it has long been recognized that standard measures of inequality – like the Gini index and others – cannot quantify the dynamics of a precarious trajectory. Indeed, precarity has been referred to as a "slow death" [94, 165] because of its progressive nature that unfolds for an individual over time.

Much research [161] has therefore gone into characterizing properties of *sequences* that describe the state of an individual over time. Researchers have proposed measures that seek to capture the *number* of distinct states, the number and direction of transitions between states, and even incorporate the significance and meaning of individual states in the sequence. For example, to capture the variability in states in a sequence, the entropy of the frequency distribution of states has been regularly used. To capture effects at different time scales, other researchers have proposed first constructing subsequences of the trajectory (akin to the use of skip n-gram models in text analysis). In this dissertation, we use one of these measures, proposed by Ritschard et al. [173], that seeks to capture three key aspects of precarity. We assume that an individual's trajectory is described as a sequence of states $\sigma = s_1, s_2, \ldots, s_t$ where $s_i \in S$ and S is the set of states. A *quality* function $r : S \to \mathbb{R}$ indicates the level of financial wherewithal (where a higher quality implies a better condition). Then the measure of precarity for a given sequence σ depends on

- The quality of the starting state $r(s_1)$
- The net decline in state over σ
- The amount of variability in σ

Net decline in state. We assume the states in S are sorted from lowest to highest "quality". In any sequence σ, we can classify transitions between states as either negative or positive, depending on which state is higher. Let $q^-(\sigma)$ be the proportion of transitions that are negative, and $q^+(\sigma)$ be the proportion that are positive, and set $q(\sigma) = q^-(\sigma) - q^+(\sigma)$. The quantity $q(\sigma)$ represents the net magnitude of negative transitions and ranges between -1 (for purely positive transitions) and 1 for purely negative transitions. We note in passing that the transitions can be weighted: in that case, the proportions are appropriately calculated in a weighted manner. We weigh them by the hops (distance) a state has to the lowest quality state in a sequence.

The variability in the sequence. There are two factors used to define variability in σ. The first is the number of states visited, or more generally the distribution of the states entered during the sequence. This can easily be captured by computing the entropy $h(\sigma)$ of the (normalized) frequency distribution of states. This in turn, must be normalized by the maximum entropy possible, which is merely $\log |S|$. This does not however capture the transitions between states. For example, consider the sequences $\sigma = (1, 1, 1, 1, 0, 0, 0, 0)$ and $\sigma' = (1, 0, 1, 0, 1, 0, 1, 0)$. Clearly $h(\sigma) = h(\sigma')$ but σ' reflects a more erratic state of existence. To account for this, Ritschard et al. [173] add in a term $t(\sigma)$ that merely counts the number of transitions to different states (normalized by $|\sigma| - 1$). Note that $t(\sigma)/(|\sigma| - 1) = 1/7$ but $t(\sigma')/(|\sigma| - 1) = 1$. These two terms are combined using their geometric mean:

$$c(\sigma) = \sqrt{\frac{h(\sigma)}{\log |S|} \frac{t(\sigma)}{|\sigma| - 1}}$$

The precarity index. The overall precarity index of a sequence $p(\sigma)$ is a function of the initial quality $r(s_1)$, the net decline in state $q(\sigma)$ and the amount of variability $c(\sigma)$. In this dissertation, we use [173]'s formulation of the index: whether other functional forms might provide different sensitivity is a matter we defer to further research. The precarity index can then be defined as:

$$p(\sigma) = \lambda r(s_1) + (1 - \lambda)c(\sigma)^{\alpha}(1 + q(\sigma))^{\gamma}$$

This can be seen as a convex combination of the starting position and terms involving dynamic components (controlled by λ). The two dynamic components are weighted by different exponents to reflect different degrees of sensitivity and importance.

We can also now elaborate on why measures like the Gini index fail to capture precarity. Precarity is a notion evaluated for an individual over time – the precarity index is a way to quantify this as a kind of time average. The Gini index instead is a measure of inequality of a population measured at a snapshot in time and acts as a population aggregate measure.

2.1.2 Fairness in Sequential Decision Making

Fairness in sequential decision-making is another pertinent piece of literature on the concept of long-term precarity in the digital age. Zhang and Liu [219, 220] present a comprehensive review of work on fairness in sequential decision making broken down by whether the decision process affects input features or not.

A large body of work considers the case where input features do not change [84, 73, 17, 97, 83, 202, 17, 120, 31, 68, 160, 46]. When considering how a population might evolve in response to decisions, two broad lines of work emerge – those that consider two decision stages [122, 85, 99] and those that consider finite or infinite-horizon decision making [89, 80, 221, 145, 123, 52].

This latter body of work is more closely related to our study. The broad goal here is to understand how qualifications of different groups evolve in the long run under various fairness interventions and the conditions to achieve social equality. They often focus on the problem of access and "dropout" (when decisions lead to withdrawal from the market) as causes of disparity between groups and propose various interventions to address this.

Another approach to understanding sequential decision making has been to take advantage of simulations on Markov decision processes (MDP). As [47] argues, long-term fairness dynamics are hard to evaluate, and so we need simulations to assess fairness over time. MDPs can also be formally analyzed for long-term effects on (group and individual fairness) as explored by [92, 96, 208, 47].

Furthermore, automated decision-making algorithms are an integral part of sequential decision-making. The general opacity of algorithms [55, 56] wherein their internal mechanics are unknown by the decision subjects results in a volatile nature which makes them prone to unexpected change. This lack of clarity can exacerbate precarity [24, 45].

2.1.3 Agent-based Simulation and Decision Making

Agent-based simulation has been used in many social settings, including fairness in lending [122], resource allocation [54, 51], college admissions [90, 100], financial analysis [32], technology adoption [106], studying supply chain shortages [215], and simulations of global crisis like the pandemic to minimize the spread of virus [2].

To study the long-term effects of decisions from the perspective of individuals (in particular, in terms of precarity), the relevant chapters in this dissertation follow in the path of research by D'Amour et al. [47] and Zheng et al. [222] which use simulation as a mechanism to study long-term behaviors of agents in systems.

2.2 Relevant Economic Models and Notions

In this section, we explain the related work on consumption models, ruin, and investment.

2.2.1 Consumption Models

The most important part of our agent-based simulation framework in Chapters 3, 4, and 5 to study phenomena such as financial precarity and work schedule instability is the process by which agents consume and earn utility. In this section, we describe the standard toolkit from the literature on consumption and then build our frameworks in Chapters 3, 4, and 5.

There are numerous models that seek to capture human consumption behavior. At their core, all of these models assume that an agent consumes an amount c in order to maximize utility $u(c)$, where $u(\cdot)$ is some concave function. Agents are assumed to receive *income* y_t at time t, as well as maintaining assets x_t. Further, most models assume some form of *discounting*: that an agent prefers to receive utility now rather than later. This is formalized by saying that the actual utility gained by consuming c_t at some future time t is $\beta^t u(c_t)$, where $0 < \beta < 1$ is a *discounting factor*.

In all models, the goal is then to determine how an agent might choose consumption c_t at each time t to maximize their long-term utility, given by $\sum_{t=0}^{\infty} \beta^t u(c_t)$. In other words, intertemporal consumption models ask the individual to perform a discounted consumption utility maximization with instant rewards being preferred to future rewards of the same size. [30]. The individual does this via choices of when to consume or save [177]

The intertemporal consumption models used widely are the permanent income hypothesis, the life-cycle model [37, 62, 159], and the neoclassical consumption model [27]. In the permanent income hypothesis (PIH) model, the agent looks into the future and calculates the amount to consume at a time point by considering the expected average income over time [61]. The life-cycle model of consumption follows a similar construction, but incorporates the idea that an individual has a certain time frame (in contrast to PIH where the agent lives indefinitely) over which they acquire assets, and the goal in this model is to maximize the gain over the assets in the given time frame.

The neoclassical model is a simplified version of the life-cycle model in which the individual considers the present time point and the future (as two time points) and comes up with a consumption for the present day and the future based on the current and the future income parameters [37, 95].

Following up on §2.1, there are two limitations in existing consumption models based on discounted utility when it comes to uncertainty and precarity. Firstly, both the PIH and life-cycle models use a weighted aggregate of income to calculate permissible consumption. As a consequence, it is possible for agents to consume (borrow) more than their available assets at any time step, which prevents us from detecting bankruptcy.

Secondly, PIH does not handle uncertainty well, and when it attempts to do so, it forces the individuals to engage in overly risk-averse savings, which does not match human behavior [206]. The neoclassical model assumes complete and perfect information about the future, and therefore, the choice of consumption is uncertainty-free with perfect rationality [11]. More generally, precarity is a measure of an individual's financial trajectory, and thus cannot be properly modeled by measures that smooth out financial behavior over time via aggregation.

The income fluctuation problem (IFP) is another relevant consumption model with dynamic optimization over an infinite time horizon [127, 180, 35, 39, 107, 166, 171, 184]. It models uncertainty in income at each time point and limits consumption to the available assets, rather than allowing for unrestricted borrowing as in the earlier models.

We also note that while utility maximization is not necessarily the only way to model agents of bounded rationality [15], it is the framework with the most extensive machinery and tools for modeling.

2.2.2 Ruin Theory and Minimum Subsistence

Ruin theory has traditionally been used to model an insurer's liability to insolvency [185, 8]. The theory has also been used to examine income shocks and optimal stimulus allocation to minimize bankruptcy [3, 158].

The idea of lower bounds on consumption – minimum subsistence – has been studied in the context of determining utility-maximizing consumption [219, 223, 225, 193, 192, 7, 46, 138, 137]. Minimum subsistence captures the notion that individuals have to consume enough to satisfy their basic needs such as food, shelter, and clothing.

2.2.3 Investment Models

The body of work on investment models looks at an investor's consumption and investment decisions in continuous time to maximize utility [104]. These decisions often introduce constraints in the form of upper bounds on the probability of going to ruin before the time of death [16, 71]. There are also works in the economics literature that introduce borrowing [109, 64], debt [136], the effects of debt on the parameters at a macro level, and controlling borrowing using constraints [138, 102, 75]. Additionally, in the field of investment, where high volatility and real-time information availability are prevalent, there is a demand for an online model that can allocate investments among a set of assets and maximize cumulative wealth through sequential optimizations. This application of online learning is commonly known as online portfolio selection [42, 214, 118, 119]. It represents an algorithmic trading strategy in online learning, wherein future prices of risky assets are predicted using historical price information. Subsequently, online learning algorithms optimize the portfolio by employing loss functions tailored to specific financial objectives, ultimately aiming to achieve maximum wealth.

While investment models [16, 71] have valuable components, including modeling of uncertainty, they can only imperfectly model consumption. Investment models involve strategic decisions about a collection of financial investments (i.e., portfolio allocation) between risky and riskless investments rather than decisions about consumption. An individual can control their chances of poverty or financial ruin in these models through the reallocation of resources, which has no analog in managing day-to-day decisions on how much to save or consume.

2.3 Temporal Instability and Lookahead

In this section, we include the related work on work timetable instability and the research in reinforcement learning on capturing the notion of lookahead.

2.3.1 Work Schedule Instability

In terms of work schedule instability, the current focus of research primarily lies in the field of sociology, specifically examining irregular work scheduling and its various repercussions. Unstable schedules cause income volatility [77, 141, 57, 172, 186] and income volatility results in financial and life hardship [13, 172, 116, 133, 111]. This encompasses issues such as burnout from precarious work schedules [187, 82] and work-family conflicts [70, 86], particularly affecting parents with unpredictable or just-in-time schedules. The impact also extends to areas like anxiety and child behavioral problems linked to parental work instability [188]. Additionally, the field of Human-Computer Interaction (HCI) has also strived to study similar repercussions with a participatory outlook [199, 212, 115].

Further, statistical data shows a pronounced unfairness in advance notices for altering work timetables for underprivileged groups. Hourly workers, individuals with lower educational attainment, women of color, and specific service sectors are disproportionately affected [188, 134] with managerial discretion [211, 110]. These work schedules make it difficult to plan for the future [191] and difficulty in planning would result in disproportionate financial poverty and hardship [65].

Along the same lines, there are some reports pointing to scheduling software and planning algorithms as a factor behind more unpredictable scheduling, particularly for low-wage workers in the service industry [101, 112, 72, 218]. For example, a New York Times article pointed out how some employees with algorithmic schedules rarely learned their timetables more than three days before the start point of a workweek [101] or how pay reduction is correlated with sudden schedule changes and sales figures [125].

2.3.2 Reinforcement Learning and Lookahead

The concept of lookahead has garnered significant attention in the recent landscape of reinforcement learning (RL). Generally, in RL models, recent works look into the idea of H-step lookahead. In H-step lookahead [194], the learner has a learned dynamic model and that is used to calculate the action sequence to a horizon of size H to find the optimal policy that maximizes the cumulative result. There are also approaches where the learner incorporates a greedy real-time dynamic programming algorithm, replacing the greedy step with an H-step lookahead policy [49]. The works of [139, 194, 33, 48] study the properties of H-step lookahead where the horizon is of a fixed size.

2.4 Feedback in Societal-level Dynamics

Our exploration of societal feedback in social systems is situated within the broader context of studying the long-term effects of fairness in the presence of feedback. §2.1.2 discusses more on the broader framework of sequential decision-making.

2.4.1 Long-term Fairness and System Dynamics

For a discussion of long-term fairness that does not explicitly model agent behavior and considers system dynamics, we consider the work by Mouzannar et al. on affirmative action [144]. In this work, group outcomes are considered under different affirmative action policies within different systems to explore the cases in which affirmative action is an appropriate policy to reach long-term equality.

For a more 'model-free' approach, we turn to the effect of feedback in the context of predictive policing [54]. The interaction between predictive policing software and policing itself are analyzed using a discrete urn model, and the feedback is shown to be positive, i.e., resulting in divergence; Police end up vastly over-policing one neighborhood, regardless of the neighborhood crime rates [54].

2.4.2 Economic Models of Inequality

Most of the economic literature on inequality is about the relationship between growth and inequality. The literature refers to either political economy or wealth effect arguments [12] in which the economy is populated by a continuum of agents who are evolving over time (using agent-behavior modeling) to either maximize individual gain or to bring about economic growth (as explained in the earlier sections of this chapter). In addition, many such studies of inequality are built upon wealth distributions where some form of general-equilibrium or quantitative models with heterogeneous agents are in place [28, 12, 93]. Other models to forecast economic inequality require a concrete understanding of the macroeconomic explanatory parameters of the system. The model requires explanatory parameters to fit historical data and forecast future inequality. Examples of such parameters include human capital attainment, labor force indicators and macroeconomic indicators, e.g., GDP and inflation [69]. Note that the Lorenz Curves [126], the Gini coefficient [142], and Theil index [140] are some of the most well-known inequality measures, but are not models that can be used to predict the trajectory of future inequities.

In a societal-level approach, we do not need to have such detailed information about the macroeconomic and explanatory parameters of the system (which might not even be available or extractable from the data). In addition, our viewpoint is broader than individual-based optimization, allowing forecasting of the production and long-term sustainability of equity in a social system.

The area of "systems dynamics" applies feedback modeling to study the complex dynamical behaviors of economic and social systems, for example, the interaction between road construction, recycling, and mining [130]. Specific feedback mechanisms, including delays, differential and/or integral effects, are assumed to exist, and specified with each model. In model-building in social work, community engagement can be used to elucidate all of the possible feedback loops in the system [88].

Chapter 3

An Individual-level Framework to Study Precarity

Adapted from: Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana, Neal Patwari, and Suresh Venkatasubramanian. "Precarity: Modeling the Long Term Effects of Compounded Decisions on Individual Instability." In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, pp. 199-208. 2021.

The analysis of the societal ramifications of automated decision-making has predominantly concentrated on assessing fairness at the point of decision, scrutinizing the equity of decision sequences or pipelines with regard to a specific population, or investigating the interactive dynamics between the decision-maker and the subject of the decision [219].

What has been overlooked in this examination is an exploration of "precarity", a term introduced by Judith Butler [25] to delineate an unstable state of existence characterized by the susceptibility of negative decisions to have far-reaching repercussions on an individual's well-being. Such ramifications extend beyond mere alterations in income or wealth or the impact of isolated decisions.

To delve into the concept of precarity, there is a necessity to shift our perspective away from the decision-maker and toward the individual subject to the decision. This entails moving away from analyzing aggregates of decisions across a population and instead focusing on the cumulative effects of decisions made for an individual over time.

An individual who lives with higher precarity is more affected and less able to recover by the same negative decision than another with low precarity. Thus including only the direct impact of a single decision or a few decisions is insufficient to judge if that system was fair. However, precarity is not an attribute of an individual; it is a result of being subject to greater risks and fewer supports, in addition to starting off at a less secure position. Precarity is impacted by racism, sexism, ableism, heterosexism, and other systems of oppression, and an individual's intersectional identity may put one at greater risk in society, subject to a lower income for the same job, less able to build wealth even at the same income level, and less able to recover from harm.

Given that automated decision systems and public policy rules operate in a world in which some people's long term well-being is impacted more by the same action, how do we account for the effects of automated decisions and, more generally, proposed public policy rules? One may advocate for pilot studies, in which the policy or algorithm is deployed on some group. However, since precarity is a long-term consequence, a pilot study will necessarily take a long time to evaluate its effects. When a policy is needed for urgent circumstances, such as addressing the impact of a pandemic, there is little opportunity for testing policies in pilots.

Thus, in this chapter, we propose a modeling framework to simulate the effects of compounded decisions on an individual over time, incorporating a quantification of their precarity. Our framework allows us to explore the effects of different kinds of decision-making processes on individuals' levels of precarity. In particular, we are able to demonstrate the ill-effects of compounded decision making on the fairness of automated decisions and policies.

While our model does not capture the full extent of the realities which place some individuals in the precarious position of being more harmed by the same decision compared to someone in a less precarious state, our model does add sufficient complexity to demonstrate how this can happen, and further, a method to quantify the effect. The message for fairness advocates is that one must look beyond the effect of a single decision on a large number of people, to look at how aggregates of decisions over time impact individuals as a function of their precarity.

3.1 Contributions

The main contributions of this chapter can be summarized as follows:

- We introduce the idea of *precarity* to the world of automated decision making, drawing on an extensive literature in sociology and economics.
- We build a simulation framework to experiment with and understand the evolution of precarity in a population. This framework incorporates ideas from macroeconomics as well as the framework of bounded rationality to capture the way income *shocks* affect the long-term dynamics of individual wealth.

- We present a suite of insights drawn from our simulation platform that validates some of the observations on precarity we see particularly visible in the context of the encountered COVID-19 pandemic and illustrates how we can evaluate the effectiveness of proposed policy interventions.

3.2 A Simulation-based Methodology for Exploring Precarity

Continuing in the line of works like [47] and [222], we use a simulation framework to explore the dynamics of precarity. In this simulation framework, individual *agents* make choices (and are subject to decisions) within a system, and are described by parameters for income, wealth and health. We use population-level economic data to initialize the system, and allow the agents to make either locally reasonable decisions (in a *bounded rationality*-like framework) or allow them to maximize expected utility within epochs. Using a simulation framework with realistic input parameters and controls allows us to observe the evolution of the system in a way that would be difficult to do formally (like for example, [3] is able to do for the more specific problem of income shocks), and allows us to experiment with different kinds of interventions.

Agents. The agents are households who interact with simulated environments in an alternating loop. Each agent is specified by their **income**, **net worth**, and **health**. An agent incurs **expenses** each time period and also earns income. Agents must make decisions about their assets – whether to consume, pay for expenses, save, or improve their health.

States. We associate each agent with a set of three states (one for each of income, net worth, and health). Each state indicates which decile of the overall population they are in for that attribute (so there are a total of $10 \times 10 \times 10 = 1000$ possible states).

Metrics. We use sequences of states for each attribute separately to perform precarity computations for each agent as described in §2.1.1. We set the needed values to calculate precarity as $\lambda = 0.2$, $\alpha = 1$, and $\gamma = 1.2$, as is done in [173]. Note that we use the term $1 + q(\sigma)$ to yield a term between 0 and 2: if the trajectory of the sequence is purely positive (thus setting $q(\sigma) = -1$) the precarity is merely a function of the initial state. In our experiments in this chapter, we test several values of α and γ, and they do not affect the results as long as they are above zero, since the overall effects on the underlying population will be similar for all data points.

We record the precarity value of all households for each income decile.

Initialization and Updates. We initialize a population of agents using parameters drawn from published statistics. For initial income, we generate an income distribution of 10,000 points using 2019 income data of the US Census Bureau's Annual ASEC survey of the Consumer Price Index (2019 dollar values) as detailed by the IPUMS Consumer Price Survey [58, 163]. To each household, we assign a net-worth (their financial and non-financial assets minus their liabilities). The net worth is assigned by detailed median percentile net worth data and median net worth by income by percentile data from the Federal Reserve.[1] The health index average of the population is extracted from the Census Bureau CPS Annual Social and Economic (March) Supplement 2019.[2] Note that we consider one health feature for the entire household. While health is of course a personal state, this allows us to combine this data with the household-based data for the other attributes. Each household has a set of basic expenditures each month (e.g., for food, housing, transportation, etc.). These expenses are extracted from 2019 mean annual expenditures from the Consumer Expenditure Surveys of the US Bureau Of Labor Statistics.[3]

Updating Health Information. Net worth automatically updates as agents spend their income and/or save it. Income updates happen via a decision process that we describe below. What remains is to describe how the health status updates. The relationship between health and income has been observed to be positive and concave [164]. Wagstaff and van Doorslaer [204] have proposed modeling this as a second degree polynomial whose first gradient is positive and the second gradient is negative. To define the function, we use the principle of *relative income theory* where "health depends on income relative to average incomes of one or more reference groups" [36]. That is, the individual (we consider the household one entity) health equals income relative to a specific group's income reduced by the square of income relative to the group's income:

$$h_i = \bar{h} + \eta(w_i - w_g) - \sigma(w_i - w_g)^2,$$

where h_i is the individual (household's) health, $\bar{h}$ is the mean health index of the whole population (extracted from CPS), w_i is individual (household)'s income, and w_g is the income mean in the group of reference, i.e., the income decile the household is in, in each round of decision making. η and σ are positive model parameters. We choose parameters that result in a wider range of indices for precarity states. We choose 1 and 10^{-20} for η and σ, respectively.

[1] https://www.federalreserve.gov/
[2] https://www.census.gov/programs-surveys/cps.html
[3] https://www.bls.gov/cex/tables.htm

3.2.1 Income Shocks

The decisions are made for 10 rounds on a monthly income and expenditure monetary value basis. The effects of positive or negative decisions are reflected on income after each round. We deduct (add) a unit based on negative (positive) outcomes if the households do not stay in the same state. We set this unit as 10% of their income, which has less financial calamity than setting a fixed value (e.g., $500) for lower incomes since it decreases proportionally with their income. Clearly, a higher value will be more beneficial for the wealthy and more ruinous for middle and lower income households.

Benefit Decision Policy. Public policy can improve household financial stability by providing benefits, and in these simulations, we explore the effect of decision classifiers which make these decisions based on an individual's current state. We introduce a lenient classifier, which accepts 50% of the initial population applying for the service based on their current income. The threshold is a global fixed value for the whole population despite their previous transitions, highlighting the fact that the decision-maker is unaware of the precariousness of the household. We implemented the experiments for a range of classifier thresholds to see the precarity of the population for the most lenient classifier, the most difficult classifier, and all other classifiers in between. We chose the most lenient classifier to consider the most optimistic scenario for assigning positive decisions. The harsher classifier has more impact on precarity levels. We try to make the default simulation specification in the interest of lower income households.

3.2.2 Strategies

The final piece of the simulation is specifying how agents behave at each time step. The economics literature typically views agents as rational (discounted) utility maximizers, and an extensive literature has developed around different stochastic models under which to maximize utility. An alternative approach is to take a viewpoint of *bounded rationality*: each agent now makes realistic choices (stochastically) from a collection of options that are locally rational, but cannot perform long-range utility maximization.

We simulate agent behavior under both of these models, which we describe below.

Rational Agents and Income Fluctuations Problem (IFP)

Referring back to 2.2, in this model [181, 182, 127, 180, 35, 39, 107, 166, 171, 184], an agent finds a consumption-asset path $\{(c_t, a_t)\}$ where a_t is the assets (net worth) at point t, and c_t is the consumption at point t, with the goal of maximizing

$$\mathbb{E}\left\{\sum_{t=0}^{\infty}\beta^t u(c_t)\right\} \tag{3.1}$$

such that

$$a_{t+1} = R_{t+1}(a_t - c_t) + Y_{t+1} \;\text{ and }\; 0 \leq c_t \leq a_t \tag{3.2}$$

Where, $\beta \in (0,1)$ is the discount factor, Y_t is non-capital income (i.e., via labor), and R_t is the interest rate on savings. For simplicity, in this chapter we will disregard gains from savings by setting $R_t = 1$.

The non-capital income Y_t is controlled by an exogenous state process $z = \{Z_t\}$. As we shall see, this is how we can introduce income shocks via decision processes.

The quantity u is the utility to the household. We use the Constant Relative Risk Aversion (CRRA) [124, 205] utility

$$u(c) = \frac{c^{1-\gamma_c}}{1-\gamma_c},$$

which is a commonly used utility function in finance and economics that captures the idea that risk aversion is independent of scale. Here, risk aversion refers to an individual's inclination to prefer low uncertainty (more predictable) but lower pay results over the results with high uncertainty but higher payoffs [152]. We pick $\gamma_c = 2$ since the utility function has a $c^{(1-\gamma_c)}$ term and with a smaller value, $u(c)$ could become imaginary given that we use $u(c-b)$ where b is their monthly basic expenditures. This is to assure that they cover their basic needs in every round (if $c < b$ then there is negative utility).

A *feasible* consumption path $(a, z) \in \mathsf{S}$ is equivalent to the consumption path $\{c_t\}$. However, $\{c_t\}$ and its asset path $\{a_t\}$ must satisfy the following:

- $(a_0, z_0) = (a, z)$
- the feasibility constraints in 3.2
- being measurable. This means that only before t (and not afterward) the consumption path is a function of random variables. Thus, at time t the consumption cannot be a function of unobserved outcomes.

An *optimal consumption path* (a, z) is a feasible consumption path that attains the supremum in max (.) for the objective 3.1, which can be shown to be:

$$u'(c_t) = \max\left\{\beta R\,\mathbb{E}_t u'(c_{t+1})\,,\ u'(a_t)\right\} \tag{3.3}$$

This can be optimized to find the optimal consumption. Please see Appendix A.2 for details.[4]

Markov Decision Process (MDP) Model

We now turn to our second approach to modeling agents. Here, each agent will occupy a state of a Markov decision process, with transitions out of each node based on locally reasonable decisions about asset management. Agents can use their savings to pay for their necessities and liabilities, they can sell a tangible asset, opt for the conversion of a health-related tangible asset (such as an insurance plan). They can use their income to increase their savings, invest in health improvement, or build assets through consumption. The decision they make (stochastically) moves them to a new state with modified attributes (income, wealth and health) accordingly. See Appendix A.1 for details on the transitions in this system. We note that the transitions are designed based on prior studies of income and precarity [147] that describe typical behaviors of individuals in different income classes when faced with income and health shocks.

The Decision-making Process. The IFP model presents challenges for the income shock process that we described in §3.2.1. In the macroeconomic literature on income fluctuation (and indeed also in the work by [3]), shocks are assumed to present in a stochastic form. Thus, while there is randomness in the shock generation process, it is a predictable kind that can be optimized for (in an expected sense). However, shocks generated by an external decision cannot be optimized for in the same way (and indeed, this is an important element of precarity). In our simulation, we think of the optimization process as happening in epochs *between* decision points. This model captures the idea that long-range planning is constrained by decision points that the individual has no control over.

3.3 An Empirical Inquiry

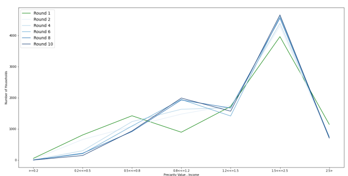
With our simulation framework now in place, we are ready to explore a set of questions relating to how precarity manifests itself. For each of these questions, we will run both simulation methodologies described above in §3.2.2. We will run the simulation for a fixed number of time steps, recording the (cumulative) precarity indices of individuals for each of the three state variables as their state

[4] Our explanations and implementation for this model are built upon `https://python.quantecon.org/ifp_advanced.html`

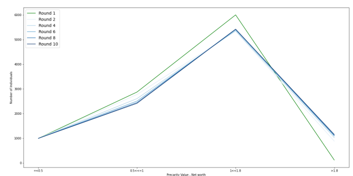
string gets longer. We will show the distribution of precarity index values across the population at each time step in order to illustrate how the distribution evolves over time.

3.3.1 Evolution of Precarity

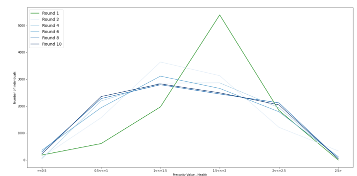
Our first sequence of experiments acts as a baseline to demonstrate how income shocks affect the precarity of individuals over time with respect to each of the three state variables. The results (for each of the variables) for the IFP model are shown in Figure 3.1 and the corresponding results for the MDP model are shown in Figure 3.2.



(a) Income precarity

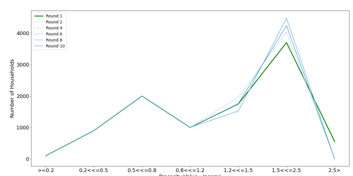


(b) Net worth precarity

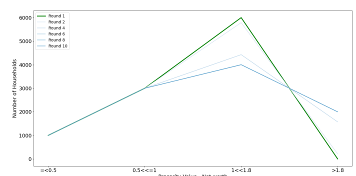


(c) Health index precarity

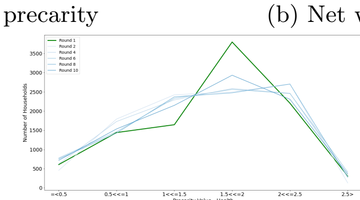
Figure 3.1: Assessing household precarity over time - IFP model



(a) Income precarity



(b) Net worth precarity



(c) Health index precarity

Figure 3.2: Assessing household precarity over time - MDP Model

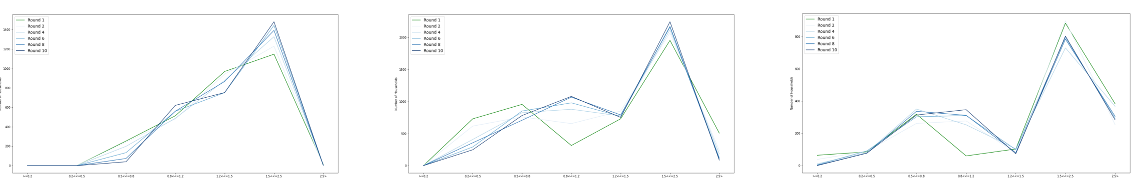
Analysis. In both models, we observe that as the system progresses the precarity distribution for net worth shifts rightward (i.e., there is an overall increase in precarity). The changes are of different magnitudes (and we will explore the reasons for that next), but it is worth noting that *income* shocks affect both net worth and health indices because of the interconnected nature of these attributes in reality.

The health index precarity changes in a less consistent manner: indeed in the IFP model it appears that health precarity appears to decrease in certain parts of the distribution. We suspect this is because of two factors: firstly, the health index is computed relative to the average income level in a particular state. Thus, even if income decreases, the health index might appear to be “further” from that mean value and spuriously indicate a better health index (see the discussion in §3.2). A second cause of this effect could also be that individuals starting off with high precarity might have their precarity *reduce* as they see similar states (even if they are inferior states): this is linked to the way in which the different terms in the precarity index are weighted.

3.3.2 Heterogeneity in Precarity Evolution

The above picture is a global view on precarity across all income levels. One of the observed effects of precarity is the non-uniform way in which individuals at different income levels might be affected by financial shocks. To investigate this, we look at precarity distributions segmented by income level. These classes are the lower 29% of the incomes, the middle 52% of incomes, and the upper 19% incomes.[5] In the IFP model, the precarity index of income, net worth, and health can be seen in Figures 3.3, 3.4, and 3.5, respectively. Figures 3.6, 3.7 and 3.8 show this for the MDP model.

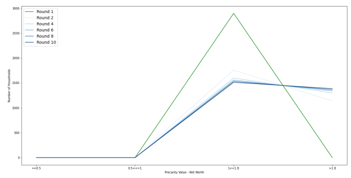
Analysis. In general, we see the following consistent behavior. Higher income individuals maintain a (low) level of precarity over time and sometimes even experience a *decrease* in precarity. Lower income individuals experience a clear increase in precarity, and middle income individuals also experience a precarity increase (but less). In other words, there is a *compounding* effect of income shocks for individuals who are already in precarious positions, the exact concern that precarity seeks to capture. While our simulation is a gross oversimplification of reality, this phenomenon has been observed in the real world. During the pandemic for example, in March around 34.4% of low income people with income less than $27,000 lost their job compared to that of only 13.2% high income people with income more than $60,000.[6]



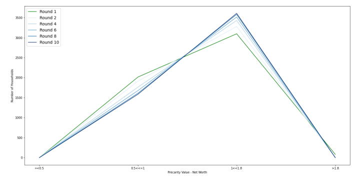
(a) Low income precarity (b) Middle income precarity (c) High income precarity

Figure 3.3: Assessing income classes' precarity over time - IFP model

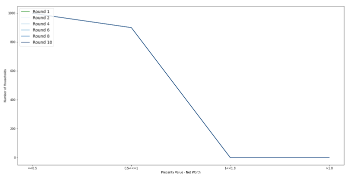
[5]`https://www.pewresearch.org/fact-tank/2020/07/23/are-you-in-the-american-middle-class/`
[6]`https://tracktherecovery.org/`



(a) Low income net worth precarity

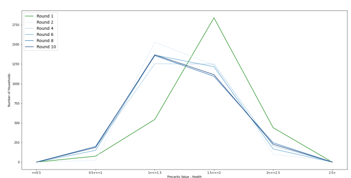


(b) Middle income net worth precarity

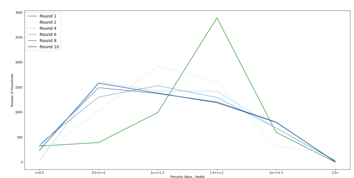


(c) High income class net worth precarity

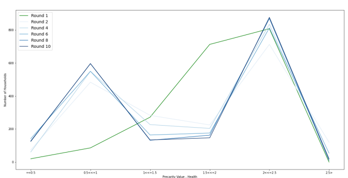
Figure 3.4: Assessing income classes' net worth precarity over time - IFP model



(a) Low income health precarity

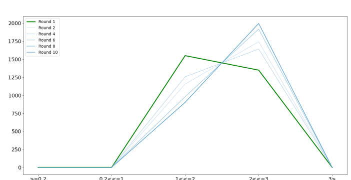


(b) Middle income health precarity

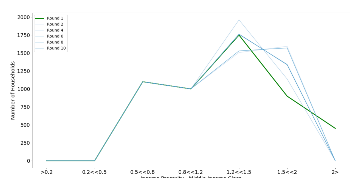


(c) High income class health precarity

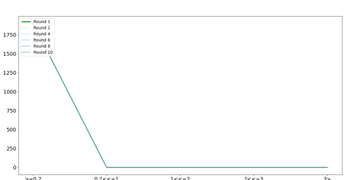
Figure 3.5: Assessing income classes' health precarity over time - IFP Model



(a) Low income precarity

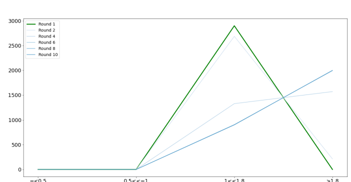


(b) Middle income precarity

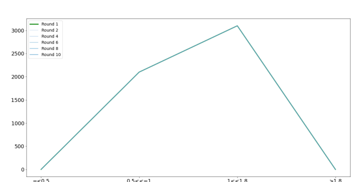


(c) High income precarity

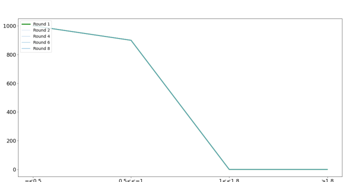
Figure 3.6: Assessing income classes' precarity over time - MDP Model



(a) Low income net worth precarity

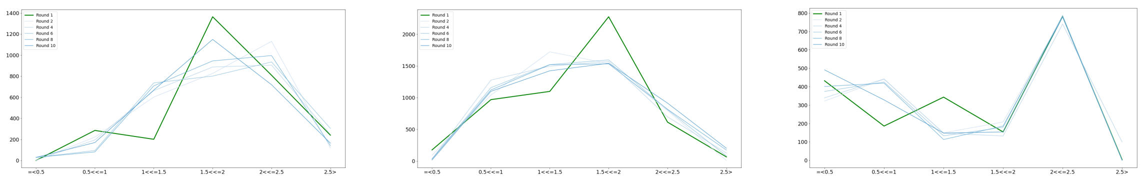


(b) Middle income net worth precarity



(c) High income net worth precarity

Figure 3.7: Assessing income classes' net worth precarity over time - MDP Model



(a) Low income health precarity (b) Middle income health precarity (c) High income health precarity

Figure 3.8: Assessing income classes' health precarity over time - MDP Model

3.3.3 Policy Interventions

A potential value of a simulation framework is our ability to experiment with interventions that would be difficult if not impossible to test "in the wild". We demonstrate the value of our simulation with two policy interventions that might be implemented to alleviate precarity. Both interventions are motivated by concrete measures that have been proposed to alleviate wealth shocks experienced during the pandemic.

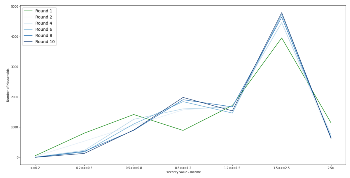
1. *Fixed stimulus intervention*: We consider a fixed stimulus intervention (measured as a fixed monthly value of \$1500 similar to the stimulus monthly checks during the pandemic [7]) given to all households who fall below the classifier threshold on every round. This form of fixed stimulus is similar to the mitigation model suggested by Abebe et al. [3] (although in their model the goal is to allocate different fixed amounts of stimulus to different individuals)

2. *Precarity resistance*: An alternate approach to dealing with income shocks was demonstrated (among others) by Germany, where the government instituted a program to help people keep their jobs and continue to be on the payroll [91].[8] We modeled this by reducing the probability of a transition to a poorer economic state after an adverse decision in our simulations. We implement this in the MDP model by adjusting the transition probabilities directly and in the IFP model by adjusting the transition process that generates the exogenous state Z.

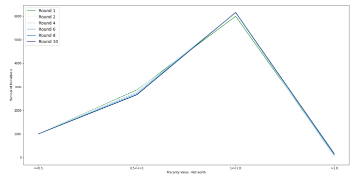
We show the out-turn of the same policy interventions in the IFP model. These results are shown in Figures 3.9 and 3.11. Figures 3.10 and 3.12 show the results for the MDP model, respectively.

Analysis. We see that these interventions have a measurable effect on decreasing household precarity compared to Figure 3.2, as the number of households with higher precarity indices reduces. The effect of enforcing an intervention on income is of a ripple effect on other tied features: we also observe a decline in precarity for these features. In addition, we see a measurable effect of the

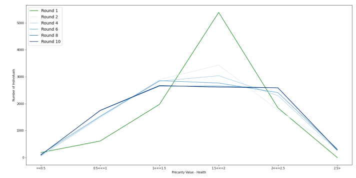
[7] `https://www.nytimes.com/article/coronavirus-stimulus-package-questions-answers.html`
[8] Germany's Kurzarbeit Program: `https://tinyurl.com/yd9qpahs`



(a) Stimulus effect on income

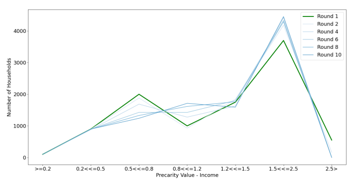


(b) Stimulus effect on net worth

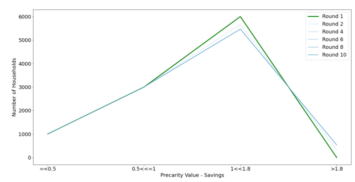


(c) Stimulus effect on health

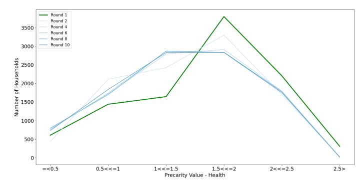
Figure 3.9: Stimulus effects - IFP Model



(a) Stimulus effect on income

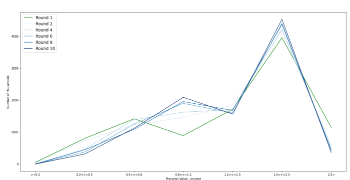


(b) Stimulus effect on net worth

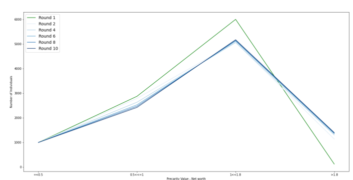


(c) Stimulus effect on health

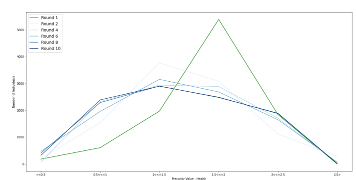
Figure 3.10: Stimulus effects - MDP Model



(a) Markov persistence effect on income

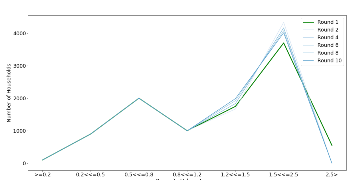


(b) Markov persistence effect on net worth

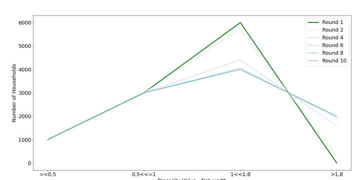


(c) Markov persistence effect on health

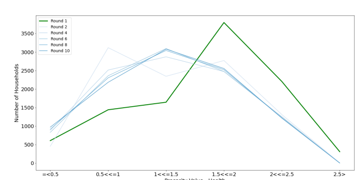
Figure 3.11: Precarity resistance - IFP Model



(a) Markov persistence effect on income



(b) Markov persistence effect on net worth



(c) Markov persistence effect on health

Figure 3.12: Precarity resistance - MDP Model

stimulus on the net worth and the count of households in the lowest income precarity values.

Note that the changes in the IFP model are more subtle because in the IFP model a household consumes as much as possible, constrained by utility and basic needs. Therefore, although assets increase with the \$1500 stimulus, it also increases the consumption. Note that we employed the interventions from the first round onward. We also tested the effects with interventions in later rounds (e.g., round 6 onward). The effects of such interventions is negligible, implying that reacting to the underlying population's precarity after they are precarious beyond the point of recovery would have little to no effect.

3.4 Chapter Summary

The main contribution of this work is the introduction of the rich sociological and the economic notion of precarity to the community of researchers thinking about automated decision making, a simulation framework to experiment with it and an empirical study of how precarity manifests in a simplified macroeconomic system.

We believe that the study of precarity is important for two reasons. Firstly, it takes the focus of decision making away from the decision *maker* and their goals for maximizing utility and other socially desirable goals, and towards the experience of an individual subject to a sequence of decisions. Secondly, this "averaging over time" reveals phenomena of inequality that are hidden underneath gross population-wide measures of progress. Lastly, although this work is noted to the Artificial Intelligence community, the contributions are also important for public policy.

In the following two chapters, we extend our investigation of this subject. Specifically, we aim to enhance the realism of the model used to explore precarity and to scrutinize the issue of precarious work environments from a temporal standpoint, focusing on the instability of work schedules. We also examine societal issues such as the US wealth gap over many decades in Chapter 6.

Chapter 4

An Agent-based Model to Study Precarity with Realistic Constraints

Adapted from: Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana*, Neal Patwari, and Suresh Venkatasubramanian. "Agent-based Simulation of Decision-making under Uncertainty to Study Financial Precarity", PAKDD 2024 (∗ represents equal contribution)*

Returning to our individual-level perspective, to study precarity – how it appears, what conditions make one precarious, and how we might mitigate it via interventions – we need to be able to model long-term financial behavior realistically rather than merely relying on existing rational models or over-simplified bounded rationality models like the MDP model in the previous chapter.

This is challenging because a) precarity is a property of an individual financial trajectory rather than an aggregate property of a population and b) it requires modeling individual behavior in response to repeated financial shocks rather than examining the effect of a single shock. Agent-based modeling [128, 32, 106, 148] is, therefore, the appropriate approach to take; if we can realistically model individual behavior, we can study how individuals respond to financial shocks and what kinds of mitigation strategies might be most effective. As an example, according to the UCSF study on homelessness in California, one finding was that a majority of homeless Californians had a median income of $960 (significantly lower than the expenses needed to maintain a house), and 70% of them believed that monthly assistance of $300-$500 would have prevented homelessness [108].

4.1 Contributions

We develop an agent-based computational model to study precarity. Our simulation framework consists of three components:

1. A mechanism for agents to maximize utility while subject to constraints under which precarity emerges – the risk of bankruptcy, consumption constraints, and uncertainty about future earnings and asset growth.

2. A mechanism to introduce income shocks into the system.

3. Mechanisms to intervene to mitigate the effects of external shocks.

This framework draws on a number of inter-related threads of work in the economics literature that seek to model human consumption, effects of uncertainty in decision-making, and challenges of avoiding ruin. One key contribution of our framework is combining all these constraints to show how an agent might maximize utility under all of these constraints.

4.2 Modeling Consumption

Like the last chapter, the most important part of our agent-based simulation framework is the process by which agents consume and earn utility. Here, we use the standard toolkit from the literature on consumption and then build our (constrained) framework in §4.3.

Capturing Uncertainty. Once again, referring back to 2.2, the *income fluctuation problem* (IFP) [127, 180, 35, 39, 107, 166, 171, 184] investigates how an agent might maximize utility when their income y_t fluctuates stochastically, while also assuming that assets earn a fixed rate of return r. To recap, the objective is to maximize:

$$\mathbb{E}\left\{\sum_{t=0}^{\infty}\beta^t u(c_t)\right\} \tag{4.1}$$

with respect to $\{c_t\}_t$ such that $x_{t+1} = r(x_t - c_t) + y_{t+1}$ and $0 \leq c_t \leq x_t$ where y_t is non-capital labor income, $\beta \in (0, 1)$ is the discount factor, and r is the interest rate on assets ($r \geq 1$). The *constant relative risk aversion* (CRRA) utility of the individual is $u(c) = \frac{c^{1-\gamma_c}}{1-\gamma_c}$, which is a standard utility function in economics [124, 205] used to capture people's preference for low uncertainty and lower gain compared to highly volatile outcomes with higher payoffs [152]. Parameter γ_c has a value $\gamma_c > 0$.

4.3 The Simulation Framework: Introducing Real Constraints

In order to model realistic agent behavior so as to capture precarity, our agents must a) try to avoid ruin (i.e., situations when the assets go below zero); b) have a fixed time horizon (i.e., a "time of death"); and c) have minimum required consumption (minimum subsistence) at each time step (for e.g., for basic necessities like food and shelter).

Formally, we define ruin as the condition $\exists t < \infty$ such that $x_t \leq 0$. Having a time of death means that the utilities are summed only till some time τ_d. Finally, minimum consumption introduces the constraint $c_t \geq c$ for all time t.

These constraints are related. For example, in the IFP model, utility-maximizing agents can avoid ruin (i.e., continue to consume into the future) as long as there are no minimum subsistence constraints (we illustrate this in a lemma in Appendix B.1). Some models avoid the issue of ruin by allowing agents to go into debt indefinitely as long as eventually all dues are paid. To account for ruin, we do not allow agents to take on debt in our model. Thus, in order to capture the above constraints, we need to make modifications to the base IFP model. To do this, we will draw on a different model of utility-maximizing consumption that arises in the context of investment.

4.3.1 Background: Modeling Ruin

Bayraktar and Young [16] introduce an investment model which has a time of death parameter τ_d, time of ruin τ_0 ($\tau_0 = \inf\{t < \infty \,|x_t \leq 0\}$), as well as a soft constraint that allows the individuals to avoid ruin before the time of death. Their model can be described as follows:

$$\begin{aligned} &\max E\left(\int_0^{\min(\tau_d,\tau_0)} \hat{\beta}^t u(c_t)dt\right) \text{such that} \\ &dx_t = [rx_t + (\mu - r)\pi_t - c_t]dt + \sigma\pi_t dZ_t \\ &\mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0) \end{aligned}$$

where x_t is the amount of current assets, π_t is the amount invested in *risky assets* like volatile stock investments, Z_t is a *Brownian motion*, $\phi(x_0)$ is a probability dependent on x_0 which is the initial amount of assets the agent starts with, and $\hat{\beta}$ is a discount factor. r is the rate of return for riskless assets (e.g., savings accounts). μ and σ are the rates for the static return and stochastic return on risky assets, respectively. Given this constrained model, Bayraktar and Young [16] construct an equivalent unconstrained optimization problem and then refer to prior work by Karatzas et al. [104] to show how this unconstrained problem can be solved. Given their dynamic equations, the unconstrained optimization problem is stated as,

$$V(x_0) = \max E\left(\int_0^{\tau_0} e^{-\gamma t}\hat{\beta}^t u(c_t)dt + Pe^{-\gamma\tau_0}\hat{\beta}^{\tau_0}\right),$$

where $P \leq 0$ is a Lagrange parameter and $E\left(e^{-\gamma\tau_0}\hat{\beta}^{\tau_0}\right)$ encodes the relaxation of $\mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0)$. The function V is called a value function, and γ comes from the time of death distribution which is exponential with parameter γ.

4.3.2 Our New Model

Recall the objective function (4.1) from IFP. It will be helpful to consider the continuous relaxation of the model. We consider a single-asset economy with a constant real interest rate. That is, let x_t, y_t, and c_t be the assets, income, and consumption, respectively, at time t and r be the real interest rate on assets. This yields a dynamic equation $dx_t = [(r-1)\,x_t - rc_t + y_t]\,dt$ and the objective, $\max\ E\left(\int_0^\infty \hat{\beta}^t u(c_t)dt\right)$ where $\hat{\beta}$ is a discount factor.

Combining this with the objective function and ruin constraints introduced by Bayraktar and Young [16] in §4.3.1, we obtain the following optimization problem:

$$\begin{aligned} &\max E_{\{c_t|t=0\dots\}}\left(\int_0^{\min(\tau_d,\tau_0)} \hat{\beta}^t u(c_t)dt\right) \\ &\text{s.t. } dx_t = [(r-1)\,x_t - rc_t + y_t]\,dt \\ &\quad \mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0) \end{aligned} \tag{4.2}$$

Note that our model is stochastic since the income decision y_t (and therefore c_t and x_t) is stochastic. Therefore, we have expectations over the sequence c_t in our optimization function, instead of just a deterministic formulation. Model (4.2) is similar in form to the Bayraktar and Young [16] model. *This key insight allows us to borrow their idea of a value function with unconstrained optimization-type formulation.* Let $c(x_t)$ be a function that returns the optimal consumption value given x_t and let $\beta = \gamma + \log(1/\hat{\beta})$. We can see that the value function still remains

$$V(x) = \max E_{\{x_t|t=0\dots\}}\left(\int_0^{\tau_0} e^{-\beta t}u(c(x_t))dt + Pe^{-\beta\tau_0}\right) \tag{4.3}$$

$$dx_t = [(r-1)\,x_t - rc_t + y_t]\,dt \tag{4.4}$$

Eliminating the time-of-ruin parameter τ_0. The optimization problem (4.3) involves a time of ruin τ_0 parameter that we do not have any information about. We need a method to remove τ_0

from the equation to derive a function that does not depend on τ_0. For this purpose, we can use a tool that was used by Karatzas et al. [104]: the Feynman-Kac formula [98]. Feynman-Kac states that, given a model $dx = \mu(x,t)dt + \sigma(x,t)dW^Q$ (where Q is a Wiener process), for any time T, the function defined $V(x,t)$ by $\frac{\partial V(x,t)}{\partial t} + \mu(x,t)\frac{\partial V(x,t)}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 V(x,t)}{\partial x^2} - g(x,t)V(x,t) + f(x,t) = 0$ can be shown to be defined by,

$$
\begin{aligned}
V(x,t) = E&\left(\int_t^T e^{-\int_t^r g(x,\tau)d\tau} f(x_r, r)dr \mid x_t = x\right) \\
&+ E\left(e^{-\int_t^T g(x,\tau)d\tau} V(x_T, T) \mid x_t = x\right)
\end{aligned}
$$

Letting $\sigma(x,t) = 0$, $\forall x$ and $\mu(x,t) = ((r-1)x_t - rc_t + y_t)$ and setting $g(x,t) = \beta$, $f(x,t) = u(c(x_t))$ and $V(X_T) = V(0) = P$ (since $T = \tau_0$ in our case) we can get the equations (4.3) and (4.4), and Feynman-Kac then gives us,

$$\beta V(x) = \frac{\partial V(x)}{\partial t} + ((r-1)x - rc(x) + y)\, V'(x) + u(c(x))$$

This gives us an optimization problem that is independent of τ_0.

Solving the optimization problem. We will now try to solve this equation. We can first remove the $\frac{\partial V(x)}{\partial t}$ to derive an equation that depends on $V'(x)$ so we have a homogeneous equation. Using the model equation (4.4) we can show that, $\frac{\partial V(x)}{\partial t} = \frac{\partial V(x)}{\partial x}\frac{\partial x}{\partial t} = ((r-1)x - rc(x) + y)\, V'(x)$. This gives us,

$$
\begin{aligned}
\beta V(x) &= u(c(x)) + 2\left((r-1)x - rc(x) + y\right) V'(x) \\
&= u(c(x)) + \frac{(r-1)x - rc(x) + y}{r} u'(c(x))
\end{aligned}
$$

where the second equation comes from solving for $V'(x)$ by taking the derivative with respect to $c(x)$ (which gives us $V'(x) = u'(c(x))/(2r)$).

Taking the derivative of $\beta V(x_t)$ with respect to x_t, we get,

$$\begin{aligned}\beta V'(x_t) - \frac{1}{2r}\beta u'(c(x_t)) &= u'(c(x_t))c'(x_t) \\ &+ \left(\frac{r-1}{r} - c'(x_t)\right) u'(c(x_t)) \\ &+ \left(\frac{r-1}{r}x_t - c(x_t)\right) u''(c(x_t))c'(x_t) \\ &+ \frac{y_t}{r}u''(c(x_t))c'(x_t) - \frac{1}{2r}\beta u'(c(x_t)) \\ &= 0\end{aligned}$$

and by setting $u(c(x_t))$ to be the CRRA utility, and using the fact that $u'(c(x_t)) = c(x_t)^{-\gamma_c}$ and $u''(c(x_t)) = -\gamma_c c(x_t)^{-1-\gamma_c}$, we get

$$\begin{aligned}0 = {} & c(x_t)^{-\gamma_c}c'(x_t) + \left(\frac{r-1}{r} - c'(x_t)\right) c(x_t)^{-\gamma_c} \\ & - \frac{1}{2r}\beta c(x_t)^{-\gamma_c} - \left(\frac{r-1}{r}x_t - c(x_t)\right) \gamma_c c(x_t)^{-1-\gamma_c}c'(x_t) \\ & - \frac{y_t}{r}\gamma_c c(x_t)^{-1-\gamma_c}c'(x_t)\end{aligned}$$

which gives us

$$c'(x_t) = \frac{\left(r-1-\frac{\beta}{2}\right) c(x_t)}{\gamma_c\left((r-1)x_t - rc(x_t) + y_t\right)} \tag{4.5}$$

Our ultimate goal was to find $c(x_t)$. We can see that given (4.5), $c(x_t)$ can be obtained by solving a differential equation. Using a symbolic solver to solve the differential equation, we get

$$x_t = k_1 c(x_t)^{\frac{\gamma_c r}{r-1-\frac{\beta}{2}}} + \frac{\gamma_c r}{\frac{\beta}{2} + (\gamma_c - 1)(r-1)}c(x_t) - \frac{y_t}{r} \tag{4.6}$$

where k_1 is a constant that we need to determine. To determine k_1, we can first use the fact that $V(0) = P$ and the fact that $\beta V(0) = u(c_0) + \frac{(y_0 - rc_0)}{r}u'(c_0)$ to derive the c_0 value which can then be plugged into the equation (4.6) to find the value of k_1. Given k_1, we have a well-defined polynomial that involves $c(x_t)$ which we can solve for any specific x_t using a polynomial solver. We defer details on how we incorporate minimum subsistence constraints into this formulation to Appendix B.2.

4.4 The Simulation Framework: Putting It All Together

We can now assemble the entire simulation framework.

Agents. Agents are initialized similar to §3.2. Agents attempt to maximize their utility in the face of an uncertain future using the approach described in the previous section. We fix the asset appreciation rate $r = 1.10$.

Shocks. A shock is a change to an agent's financial state (i.e., observable economic features which are income and assets). Shocks can affect their decisions on how to consume and save. In this chapter, we consider shocks that affect income either positively or negatively. We include two variants of real-world income shocks: *permanent* shocks and *temporary* shocks. Permanent shocks are changes to income that last until the next permanent change (e.g., unemployment, promotions, and demotions). Temporary shocks are changes that only affect the income at a specific point of time (e.g., a one-time pay cut/bonus, reduction of work hours, or economic impact payment) and do not have a lasting effect on income [103].

Algorithmic Decision Making In our experiments, we use a classifier that makes "decisions" about an agent, yielding a shock (positive or negative). Given a shock value w, if the agent gets a positive outcome from the classifier, income is changed by a multiplicative factor of $1 + w$ and if the agent gets a negative outcome from the classifier, income is changed by a multiplicative factor of $1 - w$. We use a shock value of 0.4 for permanent shocks and a shock value of 0.6 for temporary shocks. The choice of 0.4 for permanent shock size is due to the ratio between the average income of two consecutive groups being around 1.4 and the choice of 0.6 for temporary shock size was to ensure that there was a clear difference between the two types of shocks. For temporary shocks, the change to the income lasts for a single time step, while for permanent shocks, the income is changed permanently until another permanent shock happens.

In our experiments, we use a classifier that is trained to predict an agent's "financial well-being" based on their income and assets. Well-being is a subjective self-reported score (based on extensive interviews) by the Federal Reserve that accounts for how people think are doing financially (if they are "doing at least okay financially" or not) [153]. We construct a training set by assigning a label of 1 to an agent based on the probability of an individual reporting that they are financially stable (derived from the well-being score). The classifier that we use is a gradient-boosted binary classifier.

Experimental Environment. The experiments in this chapter were carried out using a Google Colab environment. The language used is Python and the experimental suite only relies on generic libraries such as SciPy, NumPy, and Matplotlib.[1]. Like before, our implementation using IFP is inspired by `https://python.quantecon.org/ifp_advanced.html`

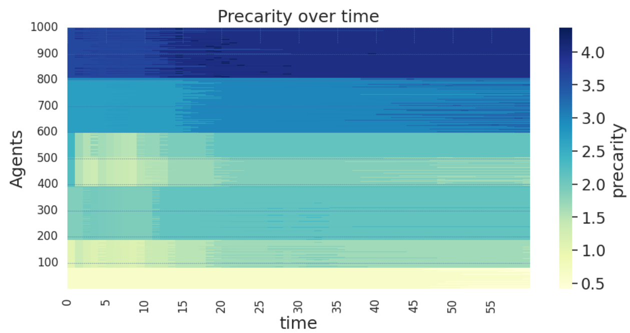


Figure 4.1: Long-term precarity analysis of a sample population of 1,000 individuals. The individuals are ordered from 1 to 1,000 based on their initial instability (lowest initial instability at the bottom). The color coding shows the precarity value and darker colors represent higher precarity.

4.5 Simulation Study: Precarity

In this section, we deploy our simulation framework to study the precarity of agents as they consume and earn in the face of financial shocks. To recap, precarity is about instability in a *sequence or history* of events rather than as a population aggregate. This is why snapshot aggregate population-level inequality measures cannot capture precarity and why researchers have developed measures of precarity based on the analysis of individual trajectories over time [161].

Precarity Index. We quantify precarity as described in §2.1.1 based on Ritschard et al. [173]'s measure, i.e., $p(\sigma) = \lambda r(s_1) + (1-\lambda)c(\sigma)^{\alpha}(1+q(\sigma))^{\gamma}$ where the initial instability is $r(s_1)$, the net decline of the sequence is $q(\sigma)$ and the amount of variability is $c(\sigma)$. The weighting hyper-parameters are once again $\lambda = 0.2$, $\alpha = 1$, and $\gamma = 1.2$ according to [173]. Note that a higher precarity index denotes higher instability and is unfavorable.

We also encode agent trajectories as a sequence of asset distribution deciles. We set the initial instability of an agent to be a function of the number of transitions required to reach the highest asset decile from their current decile and their perceived well-being (as defined in §4.4), as follows. We scale the transition-based difference by the inverse of the perceived well-being to have a more comprehensive initial instability based on both actual monetary values and the perceived self-reported economic well-being of people. Formally, for each agent, given their transition-based difference

[1] `https://scipy.org/` `https://numpy.org/` `https://matplotlib.org/`

$\Delta(s_1) = 1 + s_{\max} - s_1$ (where $s_{\max}$ is the best possible state and s_1 is the current state), and a well-being value ξ (in the range of 0 to 1), we set the initial instability to be $r(s_1) = \frac{\Delta(s_1)}{\xi}$.

As observed by Nokhiz et al. [148], the precarity measure becomes less sensitive to any change as the length of the trajectory increases. This is because of the potential decrease in the relative number of transitions as a fraction of sequence length as well as a reduction in the variability of the visited states over time. In this chapter, we mitigate this concern by evaluating precarity over a (sliding) window (of size 10); this provides a sufficiently long enough history of events to compute precarity at each time point.

4.5.1 Long Term Precarity

In this first experiment, our goal is to see how automated decisions and shocks affect long-term population precarity. We run the simulation framework as described in §4.4. For a sample population of size 1,000, we run this experiment for 60 timesteps (months). Each individual is given a lifetime of 60 timesteps with permanent shocks happening every 25 timesteps. Temporary shocks happen every 20 timesteps. The results are shown in Figure 4.1.

Analysis. There are two takeaways from the figure. Firstly, we can see the overall increase in precarity is directly linked to people's initial instability – the higher the initial instability is, the more precarity increases, culminating at an unfavorable high precarity value. That is, agents with low precarity (e.g., agents 0 to 100) have a corresponding lower precarity over time. More precarious sub-populations (e.g., agents 200 to 800), however, observe an overall increase in their long-term precarity with respect to their initial instability. Secondly, there is variability within groups with similar initial instability. This means looking at the starting condition alone is not enough to predict where someone will land in the future.

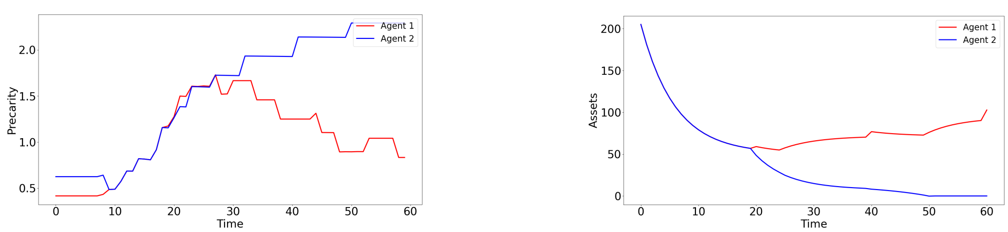


Figure 4.2: Two agents (a gig vs. office worker) with different latent initial instability, similar initial assets, and similar starting income distributions. The colors red and blue correspond to the agents 1 and 2, respectively. Assets are in thousands of dollars.

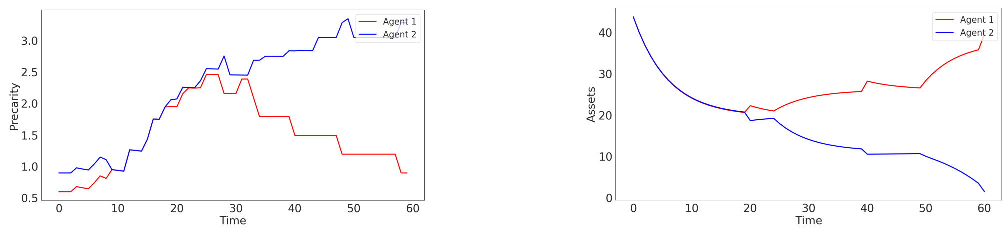


Figure 4.3: Two agents with different initial instability and initial assets ($43,800) with marginally different initial incomes. Agent 1 (*red line*) has a monthly income of $3,930 and Agent 2 (*blue line*) has an income of $3,910. Assets are in thousands of dollars.

4.5.2 Factors Contributing to Precarity

The first simulation showed the long-term effects of financial fragility, as well as how (up to a point) starting conditions can determine one's financial outcome. In this set of experiments, we go a little deeper into the hidden or latent factors of an individual's financial state (that can lead out varying outcomes) that contribute to overall financial precarity. This is important for policymakers as well as for algorithmic decision-making: if the goal of interventions is to improve financial conditions for individuals, then it is important to "see" the latent factors that can affect individual responses.

The approach we take is to compare pairs of individuals that differ in one key aspect and examine the evolution of their financial state over time. In the first scenario, the latent factor we control is income instability – the **"gig worker vs. office worker"** case. In the second, we look at individuals with marginally different incomes – **"minor income difference"**. In our plots we show results for two such individuals – we repeated these simulations for multiple pairs that fit each scenario and obtained similar results.

Our goal is to examine: 1) if hidden differences (in latent instability) can lead to largely different financial outcomes and 2) if marginal differences in observable features (portrayed in the marginal difference in initial income) can lead to vastly different aftermaths.

For the first scenario, we consider agents with two different profiles. Agent one has lower initial instability and agent two has a higher initial instability (even though their observable features i.e., assets and income, are exactly the same for an initial period). We can think of the first individual as an office worker with a stable income and the second individual as a gig worker with more latent instability and unstable income. The simulation environment is the same for both agents. After 20 months, and motivated by the volatile/insecure nature of gig work, Agent 2 experiences a drop in income. Agent 2's income distribution shifts to a lower income level (chosen uniformly at random from the range of low-income values, in our experiments).

In the second scenario, the two agents have exactly similar assets. They only have an income difference of $20. All other experimental parameters are exactly similar for both agents. As before, we simulate agent behavior using our simulation framework. Each agent is given a lifetime of 60 timesteps (months) with permanent shocks happening every 25 timesteps. Temporary shocks happen every 20 timesteps.

Analysis. There are four takeaways from this experiment. Firstly, (see Figure 4.2), an automated decision-maker that only looks at observable features, e.g., income and asset values (which are the same for both individuals) to assign snapshot decisions cannot account for diverging consequences as a result of the (hidden) instability. Secondly, in Figure 4.2, we observe the effects of the magnitude of precarity: large hidden differences in instability could lead to considerably different financial outcomes. Thirdly, in Figure 4.3, we observe the extreme difference in the precarity and asset trajectories of agents who are only marginally different in their income values. This illustrates that although their initial finances are very similar, small differences can lead to substantially differential financial outcomes. Lastly, Figure 4.3 also shows that one initial negative outcome – the initial decision – can have consequences that get amplified by the subsequent set of automated decisions made for an agent. This illustrates the way in which compounded decisions can have a significant effect on an individual trajectory.

4.6 Simulation Study: Interventions

Simulation can serve as a sandbox to test interventions that would otherwise be difficult if not impossible to explore in the real world. Fiscal interventions are actions taken by the government in the form of a collection of different subsidies, tax rebates, and unemployment benefits to address different economic circumstances. The Coronavirus Aid, Relief, and Economic Security (CARES) Act is a prime example of such a stimulus package [196]. We investigate the benefits of two types of fiscal interventions: a) tax incentives in the form of tax breaks based on the agents' income tax brackets and b) direct subsidies (e.g., COVID-19 stimulus checks).

We measure the effectiveness of interventions in terms of their *durability*. We define durability in two ways. Firstly, and temporally, we measure durability as the number of timesteps the agents would have more money compared to the scenario in which they received no interventions, i.e., their asset value be more than the baseline of the assets they would have gathered without any interventions. Secondly, and financially, we measure durability as the net difference in assets (at the end of the simulation) compared to not getting any help at all.

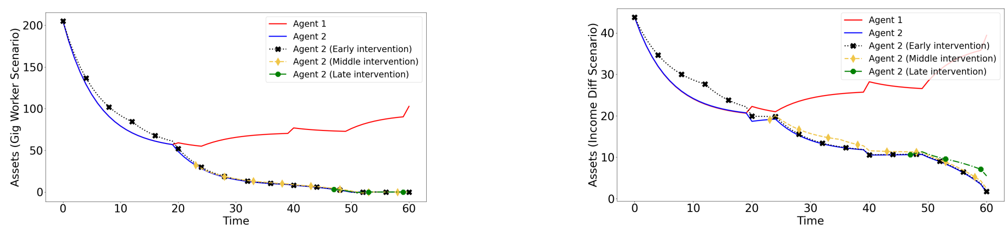


Figure 4.4: Assets with 12-month tax breaks for the exact same pair per scenario in §4.5.2, Figures 4.2 and 4.3. Here, "gig worker" is the left and "income difference" is the right figure. Markers in black, yellow, and green depict early, middle, and late interventions.

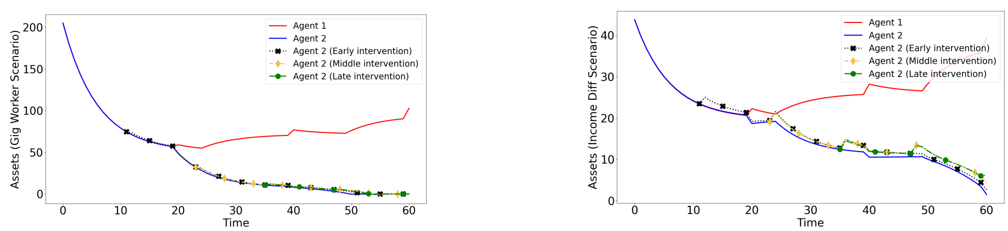


Figure 4.5: Assets with subsidies (of $3,000) for the exact same pair per scenario in §4.5.2, Figures 4.2 and 4.3. Here, "gig worker" is the left and "income difference" is the right figure. Markers in black, yellow, and green depict early, middle, and late interventions.

4.6.1 Tax Incentives

In this setting, we provide the agents with tax breaks according to their corresponding income bracket for tax season 2019-2020 [50] which effectively increases their income to the gross value for that bracket. We wish to examine three different intervention dimensions: when the intervention is made, how long the effect of the intervention lasts (i.e., the durability), and what effect changing the span of the intervention has.

As in the previous section, we compare pairs of agents who differ in one of two ways (income instability or a minor income difference). Each agent is given a lifetime of 60 timesteps (months) with permanent shocks at every 25 months. Temporary shocks happen every 20 months. We explore two lengths of 12 and 6 months of tax breaks with different start points for 50 pairs of agents in each of the two scenarios above, as follows. In an **early** intervention when agents are latently different based on precarity but have similar observable financial features, we assign the suffering second agent one-year (6 months if exploring a 6-month span) tax breaks during the first month. In a **middle** intervention, we provide the agent with a tax break at timepoint 24 for a year (or 6 months). In the **late** intervention we explore another 1-year (or 6-month) tax break starting at timepoint 48 and after several shocks and automated decisions have already occurred.

Analysis. Not surprisingly, interventions work. Durability increases for pairs of agents, as depicted (on average) in Table 4.1. Figure 4.4 is a pictorial example of this for the 12-month case for one pair – **the exact same pair** studied in §4.5.2. The corresponding after-intervention precarity plots are in Appendix B.3. There are three specific takeaways from this experiment as well. First, in general, earlier interventions when agents are latently different but have similar observable features result in improved agent durability (While there is one case when the middle intervention yields slightly better results month-wise, the difference relative to the early intervention is extremely minor in comparison to the difference with the late intervention.) Secondly, the earlier interventions help both temporal and financial durability. Lastly, the longer the intervention time span, the more durable the agents will be both in terms of assets and months.

4.6.2 Direct Subsidies

Direct subsidies, also known as stimulus checks, are direct payments to help people in financial need. In this setting, as before, our goal is to examine the three factors of the start time of the intervention, resulting durability, and dependence on the intervention amount.

Table 4.1: Mean intervention durability for 50 pairs of agents in each scenario with early, middle, and late tax breaks (studied separately for 12 months and 6 months). Assets are in thousands of dollars and rounded to the closest $1,000.

Scenario	Months			Assets		
	Early	Mid	Late	Early	Mid	Late
Gig Worker (12 m.)	**41.3**	19.2	4.7	**110**	20	7
Gig Worker (6 m.)	**33.6**	16.8	4.6	**54**	9	5
Income Diff (12 m.)	**33.4**	33.1	11.1	**64**	39	18
Income Diff (6 m.)	24.8	**25.2**	11	**32**	20	13

We explore two values of subsidies in the amounts of $3,000 and $600 (similar to the smallest subsidy

Table 4.2: Mean intervention durability for 50 pairs of agents in each scenario for early, middle, and late subsidies (studied separately with $3K and $0.6K paychecks). Assets are in thousands of dollars and rounded to the closest $1,000.

Scenario	Months			Assets		
	Early	Mid	Late	Early	Mid	Late
Gig Worker (3K)	**36.1**	25.3	12.9	**53**	45	25
Gig Worker (0.6K)	**29.4**	21.8	11.1	**12**	11	6
Income Diff (3K)	**46.1**	35.5	23.3	**47**	46	31
Income Diff (0.6K)	28.4	28.4	20.6	9	**10**	8

out of three COVID-19 installments during the two years of the pandemic peak in 2020-2021) with different start points for 50 pairs of agents in each of the “minor income difference” and “gig vs. office worker” scenarios examined above. Like before, we experiment with interventions that are *early* (3 subsidy installments every 12 months starting at timepoint 12), *middle* (3 payments every 12 months starting at timepoint 24), or *late* (3 payments every 12 months starting at timepoint 36), with two different subsidy amounts ($3,000 versus $600) in each intervention.

Analysis. As with tax relief, subsidies work, as we can see in Table 4.2. Similarly, Figure 4.5 is a pictorial example of these results for the $3,000 case for one pair – **the exact same pair** in §4.5.2. Corresponding post-intervention precarity plots are in Appendix B.3. The takeaways are similar to what we see with tax incentives. That is, knowing the diverging latent precarious nature of agents beforehand (as early as possible) can help policy-makers, in the long run, both in terms of time and money. The same amount of funds allocated for a specific type of intervention could be optimally used if it is disbursed before a diverging point in observable features when agents are only latently precarious. Moreover, bigger stimulus checks help the agents amass more assets and stay afloat longer.

Statistical Corroboration. The main goal of fiscal stimulus is to “maximize the near-term boost to economic growth without weakening the economy’s longer-term prospects. This requires that the plan be implemented quickly; that its benefits go to those hurt most by the economy’s problems; and that these benefits not damage longer-term fiscal conditions” [217, 216]. Our findings illustrate the effectiveness of early interventions when providing fiscal stimulus.

Published statistics after the CARES Act funds were disbursed show increased financial resilience to adverse shocks among lower-income households. [10]. This durability is reflected in our findings. In another statistic, payments provided under the Consolidated Appropriations Act (payments of $600 per eligible taxpayer) mostly disbursed in January 2021 resulted in personal savings increase from 13.5% in December 2020 to 20% in January 2021. This shows household wishes to save the money [60] which matches our results on saving patterns and asset accumulation after interventions.

In terms of economic stress and assets, economic stress negatively affects families’ ability to save. Economic stress resulted in less than 40% chance of household savings (according to data from 2010 to 2016), while families with no such stress had more than a 50% chance of saving [207]. This matches the overall trend we observed in the two scenarios we examined where compounded adversity and latent volatility would result in fewer assets. We note that Nokhiz et al. [148] also report on the effect of compounded shocks.

4.7 Chapter Summary

The main contribution of this chapter is a realistic agent-based simulation framework for exploring financial insecurity precipitated by algorithmic decision-making through the lens of precarity. We incorporate consumption bounds, time of death, and ruin as human-analogous constraints into our framework because the more realistic the constraints and parameters in our model are, the more we can hope to glean insights from our study. There are several ways in which we could develop this framework further, including incorporating variability in the appreciation of assets, the ability (within limits) to take loans, and adding other dimensions to the individual experience of precarity (such as factors relating to health and macroeconomic parameters like inflation).

Chapter 5

Future Lookahead: Unstable Work Schedules and Financial Security

Adapted from: Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana*, Aditya Bhaskara, and Suresh Venkatasubramanian. "Counting Hours, Counting Losses: The Toll of Unpredictable Work Schedules on Financial Security", Currently Under Submission (∗ represents equal contribution)*

In the previous chapters, we discussed how financial insecurity is becoming a widespread issue, propelled by rising financial shocks, compounded automated decisions, and inadequate wages. Continuing our examination of individual instability, this chapter also maintains a focus on the individual level.

While research on precarious work environments typically centers around financial aspects, there is a tendency to overlook the temporal aspects of unstable work schedules [188]. One cannot plan a life without a stable work schedule; therefore, the consequences extend beyond burnout and work-family conflicts [187, 82, 70, 86] and manifest as financial shocks that directly impact workers' income and assets. Unforeseen fluctuations in earnings present challenges in financial management and affect decisions on how much to save and consume. This ultimately diminishes the financial stability and long-term financial welfare of affected individuals.

This problem is particularly glaring in sectors where workers have frequently varying work schedules and receive little-to-no advance notice about changes [188, 77, 141, 57, 172, 186]. The lack of advance notice disproportionately affects vulnerable subgroups, including the food service and retail sectors, part-time and hourly workers, lower-income classes, those with lower education levels, and specific racial groups [188, 187, 134]. People belonging to these groups on average, are more financially precarious to start with, and it seems likely that the unpredictable nature of their work schedule further exacerbates their financial fragility [13, 172, 116, 133, 111].

Our goal in this work is to study this issue quantitatively. Can we understand the degree to which one's ability to manage their finances depends on one's ability to see into the future and plan? We answer this question in this chapter by building a simulation framework that models how individuals maximize utility in the presence of financial uncertainty and the need to avoid ruin. We use online learning, specifically an adaptive update of workers' consumption policies, recalculated at each step as more information about work schedules becomes available.

With this framework in place, we can show formally as well as empirically how the ability of a worker to anticipate their schedule changes enhances their long-term utility, and that conversely, the inability to see into the future can exacerbate workers' financial precarity.

This modeling approach is advantageous for several reasons. First, offline models generate a singular, fixed policy incapable of adapting to new financial information/shocks. Secondly, the simultaneous occurrence of financial data availability and optimization better aligns with real-world behavior. Thirdly, the introduction of future lookahead allows for a nuanced exploration of its utility under different conditions and parameters (i.e., different levels of foresight). Lastly, the online framework with future information provides a robust simulation environment for studying the consequences of work schedule instability and bias in future information availability.

This also opens avenues for empirically exploring mitigation strategies. This is in line with the current ongoing efforts in adopting various policies and regulations to make altering work schedules more equitable, e.g., the San Francisco Board of Supervisors' mandates (based on the Retail Workers Bill of Rights) that ensure more advance notice for hourly workers in retail chain stores when establishing or altering work schedules [70].

5.1 Contributions

Overall, the main contributions of this chapter are:

- **Online Algorithm with Lookahead:** A novel online algorithm is proposed, capable of handling varying levels of lookahead. The algorithm allows individuals to utilize the deterministic information from the lookahead to modify their consumption decisions in real-time, responding to and adapting to financial shocks as they occur. This algorithm becomes a valuable tool for studying the impact of different degrees of foresight on decision-making.

- **Formal Analysis of the Effects of Lookahead:** Workers who possess a lookahead benefit from an advantage that increases proportionally with the magnitude of their lookahead, as opposed to workers lacking any lookahead. Furthermore, it's important to note that this gap is tight under appropriate assumptions.
- **Empirical Analysis of Future Information Provision:** Individuals (workers) are provided with future information regarding certain upcoming events. This aspect enables an empirical exploration of the extent to which future information aids in financial management, contributing to increased long-term utility.
- **Temporal Equity in Workplace Schedules:** We explore temporal equity, particularly in the context of the implications of lack of advance notice (future lookahead) on work timetables that affect the disadvantaged subpopulations more acutely.
- **Mitigation Strategies:** Various mitigation strategies with reference to fair workplace laws and acts are explored to understand how the adverse effects of just-in-time work schedules, particularly in terms of workers' utility, can be alleviated. This investigation aims to identify effective strategies for improving the overall well-being of workers facing schedule instability.

5.2 Determining Consumption and Savings: An Adaptive Online Algorithm

Referring back to the concepts mentioned in §2.2, a crucial component of a framework designed to investigate the significance of advance notice (future lookahead) involves creating a model that captures consumption under lookahead while accommodating uncertainty. Specifically, it investigates how individuals, facing financial uncertainty, make decisions regarding the amount to consume and save. Although various models attempt to represent consumption under uncertainty, they all fundamentally rely on the concept of discounted utility, which is the most common model in economics for understanding the interplay between consumption and savings (as mentioned in consumption models in §2.2). To recap, in a discounted utility model, the agent, at each time step, consumes a certain amount c and receives utility $u(c)$ from some concave function u. The objective is to devise a policy to determine a consumption value c in a way that maximizes the total discounted utility. After this brief introduction, we will formally articulate the specific cases we study, assumptions, and models in the following section (§5.2.1).

5.2.1 Our Models

In this section, we introduce the main models that we study in this work (by recalling the discounted utility consumption models described in §2.2). We consider both deterministic and stochastic models that use a discount factor β to compute the utility. We assume that time is discretized into integer steps, and let T be the effective overall job timeframe (in other words, the time horizon for the algorithm). Let a_t, c_t, y_t be the assets, consumption, and income at time t, respectively. We also let R_t denote the gain from assets, i.e., the appreciation/depreciation rate of assets. Let $u(.)$ (in our context, we employ $u(c) = \sqrt{c}$ which is in the class of isoelastic utility functions that are commonly used in macroeconomics) be the utility. Let $\mathcal{D}_Y$ be the income distribution and $\mathcal{D}_R$ be the returns distribution.

In all the models, the goal of the algorithm is to maximize the total utility. More formally, it is to solve the following optimization problem:

$$\begin{aligned} \max \ & \sum_{t=1}^{T} \beta^{(t-1)} u(c_t) \\ \text{subject to:} \ & a_{t+1} = R_t(a_t - c_t) + y_t \\ & 0 \le c_t \le a_t \end{aligned}$$

The constraint $0 \le c_t \le a_t$ ensures that the worker consumes from the assets available to them. This constraint ensures the worker could consume without going to ruin. The model equation $a_{t+1} = R_t(a_t - c_t) + y_t$ shows how assets evolve over time given the consumption and income.

The offline or deterministic model. In the offline model, the income and return values y_t, R_t are known a priori to the algorithm (as, of course, is the starting asset value, a_1). This is the most common model considered in traditional economic theory, and the solution can be found using dynamic programming. The "states" in the dynamic programming simply correspond to the time step of interest and the total assets remaining.

The stochastic model. In the stochastic model, the income and return values are stochastic and they come from known distribution. In this case, an algorithm can optimize the modified objective:

$$\max \sum_{t=1}^{T} \beta^{(t-1)} \ \mathbb{E}(u(c_t))$$

The parameters and the constraints follow the same definitions as in the deterministic case, with the caveat being that y_t, R_t are updated using the "realized" values (not their expectations). Note that this is already an online algorithm: at every t, the algorithm computes the consumption value using the expected y, R for the future, but as the y_t, R_t values get updated, the algorithm changes its behavior.

Finally, to explore the effects of future information and lookahead, we consider a combination of the two models.

5.2.2 Online Consumption Algorithm with Lookahead

We finally consider a model that has "limited determinism" controlled by the extent of lookahead, and the rest of the process is stochastic. Formally, at any time t, the algorithm knows the exact values of y_{t+i} and R_{t+i}, for $i = 1, 2, \ldots, \tau$, for some lookahead parameter τ.[1] Additionally, it knows the distribution of the parameters for later time steps. Once again, in contrast to offline economic models, our online adaptive model finds an adaptive consumption policy, i.e., a policy that changes over time given the lookahead as more information arrives.

The algorithm itself is a straightforward combination of the deterministic and stochastic cases. At each step, the algorithm computes:

$$\max \sum_{t=0}^{\tau} \beta^t u(c_t) + E\left(\sum_{t=\tau+1}^{T} \beta^t u(c_t)\right)$$

Note that the algorithm incorporates new data (as well as lookahead information) as it arrives, which is why it is an online algorithm. For simplicity, we assume that a lookahead of τ implies the agent is aware of the exact return and income values for the next τ time steps, and for the time steps beyond τ, distributions of these parameters are known.

For this model, we define Algorithm 1. This algorithm solves a stochastic dynamic programming (DP) at the start (in Line 2) where it populates the table V with the maximum utility values for each asset and time point pair in the possible paths of the stochastic system. With this in hand, at time point t, the goal is to run a deterministic DP in Line 4 using the precalculated table V and use that information to calculate the optimal consumption choice at that time point t. This gives us an online algorithm that yields a new consumption at each time point t, based on the set of historical choices as well as the lookahead information available.

[1] It is also interesting to consider further hybrid models; e.g., the algorithm has a lookahead over y_t, but not over R_t. This is also realistic in practice since return rates are governed by the market while income lookahead can be controlled by the employer. We do not consider such models in this work.

Algorithm 1 Online consumption with lookahead

1: Let τ be the lookahead, x_0 be the initial assets

2: Solve the following optimization problem and save the maximum utility for each x, t in a table V where $V[x, t]$ gives the utility of x, t pair,

$$\max E\left(\sum_{t=0}^{T} \beta^t u(c_t)\right)$$
$$x_{t+1} = R_t(x_t - c_t) + y_t \text{ for all } t$$
$$y_t \in \mathcal{D}_Y, R_t \in \mathcal{D}_R \text{ for } t \in \{0, \ldots, T\}$$

3: **for** $r = 0$ to T **do**

4: Solve the following optimization problem to get c_r given x_r using dynamic programming,

$$\max \sum_{t=r}^{r+\tau} \beta^t u(c_t) + V(x_{\tau+1+r}, \tau + 1 + r)$$
$$x_{t+1} = R_t(x_t - c_t) + y_t \text{ for all } t$$
$$y_t, R_t \text{ exactly known for all } t \in \{r, r+1, \ldots, r+\tau\}$$

5: Set $x_{r+1} = R_r(x_r - c_r) + y_r$

6: **return** The sequence of c_ts at each time t

5.3 Theoretical Analysis

In this section, we establish a few theoretical results on the consumption behavior and the "power of lookahead". We analyze the advantage that a worker privileged with lookahead, could attain compared to an underprivileged worker without lookahead. Let y_t, x_t, c_t be the income, assets and consumption at time t. Let T be the timeframe of the job and $u(c) = \sqrt{c}$ be the utility gained from consuming c.

Our first result shows that even in the very simple setting of $\beta = 1$ (no discount factor), and income y_t in the range $[0, Y]$, the difference between the total utility of algorithms with a k-step lookahead and without lookahead is $\Omega(k\sqrt{Y})$. This is true not only of the algorithms we study, but *any* algorithm. This indicates that there are instances where a worker with lookahead privileges gains an edge over an underprivileged worker without lookahead, and the advantage grows linearly with the level of lookahead.

Theorem 1. *Consider two individuals, one with a lookahead of k steps and one with no lookahead. Let $c_1, c_2, \ldots, c_T$ be the consumption of the individual with lookahead k and $z_1, z_2, \ldots, z_T$ be the consumption of the individual with no lookahead. Then, there exist instances where all incomes are in the range $[0, Y]$, such that*

$$\sum_{t=1}^{T} \sqrt{c_i} - \sum_{t=1}^{T} \sqrt{z_i} \geq \Omega(k\sqrt{Y})$$

While our lower bound result is strong for large values of k, it is not very strong for small values. This is partly because we want to emphasize that a gap that grows with k holds even with $R_t = \beta = 1$.

Proof. As discussed above, we consider a very simple setting: $\beta = 1$, returns $R_t = 1$ for all t. Further we will assume that the individuals start with $a_1 = 0$ (no initial assets).

We consider the following input. Let Y be any parameter > 0.

$$y_t = \begin{cases} Y & \text{for } t \leq k/2 \\ x \cdot Y & \text{for } k/2 < t \leq k \end{cases}$$

where x is a value uniformly sampled from $[0, 1]$. Note that both the individuals (the one with and without lookahead) know this input distribution. For simplicity, we also assume that the total time horizon T equals k (this assumption can be easily removed by setting $y_t = \frac{(1+x)}{2} Y$ for all $t > k$). As a final simplification, since the incomes can all be scaled, we can assume that $Y = 1$ for the remainder of the proof.

First, consider the individual with k lookahead. Note that they can see the value of x, and thus they can consume an amount $\frac{(1+x)}{2}$ at every time step. This is feasible because the first $k/2$ steps have income $y_t = 1$, and so the assets remain above the consumption at all time steps. This yields a total utility (recalling that $T = k$) of $k\sqrt{\frac{1+x}{2}}$.

Now, consider an individual who does not have any lookahead. Intuitively, they cannot guess the value of x, and thus consuming $\frac{1+x}{2}$ is not feasible. But note that the individual may use some complex (possibly randomized) algorithm that consumes non-uniformly and achieve a high total utility. We show that this is not possible.

The starting point of the proof is the classic minmax theorem of Yao [143]: for a given input distribution, an optimal algorithm for inputs from drawn this distribution, is a deterministic algorithm. In other words, in order to prove our desired lower bound, it suffices to restrict ourselves to deterministic algorithms and prove a bound on the difference in total utility, in expectation over the choice of x. For any deterministic algorithm, since in time steps $1, \dots, (k/2)$, the algorithm only sees income of 1, the values consumption $z_1, z_2, \dots, z_{(k/2)}$ are all fixed. Let $S = z_1 + z_2 + \cdots + z_{(k/2)}$.

First, suppose it so happens that

$$\left| S - \frac{k}{2} \cdot \frac{1+x}{2} \right| > c \cdot k, \tag{5.1}$$

for some parameter c. In this case, we will argue that $\sum_i \sqrt{z_i}$ is significantly smaller than $k\sqrt{\frac{1+x}{2}}$. The starting point is the following inequality about the strict concavity of the square root function:

Lemma 2. *Let $a \in (1/2, 1)$ be a constant, and let $w \in (0, 1)$. Then we have:*

$$\sqrt{w} \leq \sqrt{a} + \frac{1}{2\sqrt{a}}(w-a) - \frac{1}{8}(w-a)^2.$$

The proof follows by a simple calculation.

Proof. We have:

$$\begin{aligned}
\sqrt{w} - \sqrt{a} - \frac{1}{2\sqrt{a}}(w-a) &= (w-a)\left(\frac{1}{\sqrt{w}+\sqrt{a}} - \frac{1}{2\sqrt{a}}\right) \\
&= \frac{(w-a)(a-w)}{2\sqrt{a}(\sqrt{w}+\sqrt{a})^2} \\
&= -\frac{(w-a)^2}{2\sqrt{a}(\sqrt{w}+\sqrt{a})^2}
\end{aligned}$$

The denominator is ≤ 8, and thus by rearranging, the inequality follows. $\square$

Now, let us write $a = \frac{1+x}{2}$. By assumption, we have that $|z_1 + z_2 + \cdots + z_{k/2} - \frac{k}{2}a| > ck$, and thus we have

$$\sum_{i \leq k} |z_i - a| > ck. \tag{5.2}$$

Next, we can use Lemma 2 to conclude that

$$\sum_{i \leq k} \sqrt{z_i} \leq k\sqrt{a} + \frac{1}{2\sqrt{a}} \sum_{i \leq k} (z_i - a) - \frac{1}{8}(z_i - a)^2.$$

Now, since the consumption cannot be larger than the overall income (which is ka), the middle term on the RHS is ≤ 0. Thus, we have

$$\sum_{i \leq k} \sqrt{z_i} \leq k\sqrt{a} - \frac{1}{8}(z_i - a)^2.$$

Next, by the Cauchy-Schwartz inequality and (5.2),

$$\sum_i (z_i - a)^2 \geq \frac{1}{k} \left(\sum_i |z_i - a| \right)^2 > c^2 k.$$

Together, the above inequalities imply that $\sum_i \sqrt{z_i} \leq k\sqrt{a} - \frac{c^2 k}{8}$. This shows that if the values z_i chosen by the deterministic algorithm satisfy (5.1), then the algorithm with no lookahead has total utility $\Omega(k)$ worse than an algorithm with lookahead.

The final step is to prove that if x is chosen at random from $(0, 1)$, the condition (5.1) holds with a constant probability for some $c > 0$. Since S is fixed, the condition is equivalent to $|\frac{2S}{k} - \frac{1+x}{2}| > 2c$. Equivalently, $|\frac{4S}{k} - 1 - x| > 4c$. For any fixed α, if $x \sim_{\text{uar}} (0, 1)$ the probability that $|\alpha - x| \leq 1/3$ is clearly $\leq 2/3$. Thus, the condition above must hold with $c = 1/12$, with probability at least $1/3$.

Putting everything together, we have that with probability $1/3$, the no-lookahead algorithm is $\Omega(k)$ worse than the algorithm with lookahead, and it can never be better. Thus the *expected* difference between the total utilities is also $\Omega(k)$. Yao's minmax theorem implies that this lower bound also holds for any (possibly randomized) algorithm. $\square$

Theorem 1 shows that an algorithm with lookahead has a provable advantage over an algorithm that knows only the distribution of the incomes, even in the simplest setting where the decay factor $\beta = 1$ and the returns are all 1. Furthermore, the advantage grows *linearly* with the amount of lookahead.

In particular, if an individual has infinite lookahead, they can have an advantage of $\Omega(T)$.

Next, we show that when $\beta = 1$ and $R_t = 1$ for all t, the lower bound from Theorem 1 cannot be improved.

Lemma 3. *Suppose $\beta = 1$ and the return $R_t = 1$ for all t. Let $y_1, y_2, \ldots, y_T$ be a sequence of income values with $y_t \geq 0$ for all t. Then if the consumption sequence of a k-lookahead algorithm is $\{c_1, c_2, \ldots, c_T\}$, there exists a no-lookahead algorithm whose total utility is $\sqrt{c_1} + \sqrt{c_2} + \cdots + \sqrt{c_{T-k}}$.*

In other words, the difference in the total utility (between the k-lookahead and the no-lookahead algorithms) is simply $\sqrt{c_{T-k+1}} + \cdots + \sqrt{c_T}$. In settings where all the c_t are of magnitude $O(\text{income})$, this corresponds to $O(k)$ times the square root of the income, which is the lower bound in Theorem 1.

Proof. The proof is simple: a no-lookahead algorithm can mimic a k-lookahead algorithm, but with a delay of k steps. We will call this the k-delay algorithm. Let $c_1, c_2, \ldots, c_T$ be the consumption squence of the given k-lookahead algorithm. The k-delay algorithm is defined as follows,

For $t = 1$ to T:

1. If $t \leq k$, consume 0
2. Else consume c_{t-k}

The total utility bound is easy to see. One only needs to check that the algorithm is feasible (i.e., it satisfies the condition that the total consumption until time t is bounded by the total income plus the assets until that time). This is easy to check because the algorithm consumes 0 for the first k steps, while the income $y_t \geq 0$. Since the decay factor $\beta = 1$, delay does not change the utility the algorithm receives. $\square$

Remark. We see that the proof relies on having $\beta = 1$. If we take into account factors such as inflation (e.g., $\beta = 0.95$), then the "power of lookahead" can likely be made much more significant.

Note on Reinforcement Learning (RL) and Lookahead. The related work on RL and lookahead has been discussed in 2.3.2. The idea of H-step lookahead in RL is similar to the lookahead employed in our models. In RL models explored in these works, there is a reliance on well-defined states and actions, involving the learning of model dynamics on the state-action space, utilizing lookahead. This learned information is then used in sampling the next state. In contrast, our models operate with no explicitly defined states.

5.4 Experiments

In this section, we explore three different sets of experiments regarding the effects of lookahead, the parameters involved in lookahead under uncertain job timetables, and mitigation schemes.

5.4.1 Experiment Setup

We will first introduce the different elements in our experiment setup.

Workers. Within our framework, agents represent employed individuals who earn weekly income, own assets, and decide whether to consume or save. We create an income distribution of 10,000 individuals using 2019 income data from the US Census Bureau's Annual ASEC survey of the Consumer Price Index (by the IPUMS Consumer Price Survey) [58, 163]. In the first stage, we eliminate outliers from the income distribution to address challenges related to individuals with exceptionally high or low earnings, which can be challenging to compare due to significant scale differences. We identify outliers by employing the commonly-used inter-quartile range (IQR) proximity rule [38]. Subsequently, we categorize individuals into four distinct income groups: low incomes ranging from $1.22 to $1125.49, low-middle incomes from $1125.49 to $2249.75, high-middle incomes from $2249.75 to $3374.02, and high incomes from $3374.02 to $4498.29. This classification is achieved by partitioning the overall income range into these four segments. To establish asset values, individuals are assigned the median population asset value of $123,840, derived from median percentile net worth data and median net worth by income percentile data from the Federal Reserve [154].

Minimum Subsistence. Furthermore, our simulation considers minimum subsistence, i.e., the constraint that the individuals must allocate funds for their minimum basic needs, such as food and shelter [219, 223, 225, 193, 192, 7, 46, 138, 137]. This consideration acknowledges the fundamental requirement for individuals to fulfill their basic necessities as part of the decision-making process regarding consumption and utility maximization. The minimum subsistence values are derived from mean annual expenditures in 2019 from the Consumer Expenditure Surveys of the US Bureau Of Labor Statistics [155], stratified based on income levels. In essence, individuals with similar income levels are obligated to cover equivalent amounts for their basic needs.

Shocks. Shocks, defined as alterations to an agent's financial state due to schedule instability, play a pivotal role in influencing decisions related to consumption and savings. These shocks can either positively or negatively impact income, such as sudden work-hour reductions or increases. The magnitude of income shocks ranges from -0.4 to 0.4, with shocks occurring as $(1 + r) \times$ income, where r represents the shock value. The shocks are generated from a Bernoulli process, with the shock size parameter r uniformly sampled from $[-0.4, 0.4]$.

Isolating the Impact of Lookahead and Other Parameters. At each time point (here, a workweek), each individual within an income group shares identical observable features, such as income, shocks stemming from unpredictable schedules, and returns on their assets. The sole point of divergence among individuals within the same group is the extent of lookahead they possess. This deliberate design choice aims to isolate the temporal impact of lookahead on utility, eliminating other financial factors' interference. The overall job timeline spans 26 weeks, equivalent to a 6-month job duration. The return range on saved assets varies between $0.9 - 1.1$ with an added variance of ± 0.05. The discounting factor β is set to the commonly-used value of 0.95.

5.4.2 Future Lookahead: An Empirical Inquiry

This section aims to examine the utility acquired by individuals under varying levels of lookahead. Specifically, the experiment seeks to compare those with minimal (or no) information about the future, such as protected groups like hourly, part-time workers, and specific racial groups that typically receive limited or no advance notice, against other workers with more foresight. The experiment setup is as explained in §5.4.1 and the outcomes of this experiment are depicted in Figure 5.1.

Analysis. The insights from Figure 5.1 can be summarized in four key points.

- Firstly, workers with lower lookahead and minimal information about future instability experience significantly lower financial utility compared to those with higher lookahead. Unsurprisingly, individuals struggle to manage their finances effectively when confronted with unforeseen schedules.
- Secondly, lookahead has a positive impact overall. Workers with more lookahead can efficiently manage their finances, resulting in higher utility.
- Thirdly, workers do not require full information about their work schedules in advance. Even beyond the midpoint lookahead (lookahead 13), workers can achieve a utility comparable to those with complete information about their schedules.
- Lastly, higher income leads to greater utility, as individuals can consume without concerns. This is evident from the consistent shift of the plots along the y-axis.

However, irrespective of income level, having more than the midpoint level of lookahead proves advantageous for individuals. Providing people with advance notice of their schedules can be

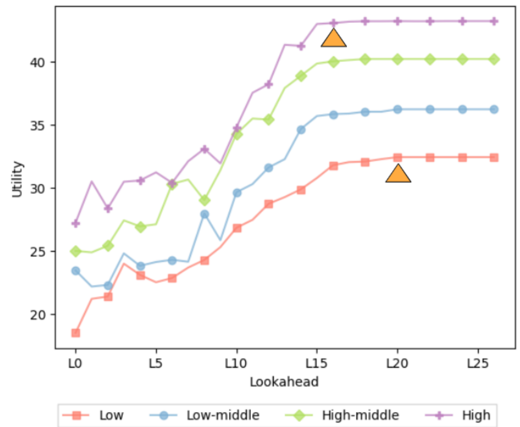


Figure 5.1: The final utility gained for different levels of lookahead is illustrated for four distinct income classes, each comprising 27 individuals. Workers are of similar features, with variations solely based on the temporal aspect, i.e., the amount of lookahead in their work schedules. The x-axis depicts the lookahead value and the y-axis represents the total utility at the end of T steps. The two orange triangles are in place to highlight the difference in the level of lookahead needed for approaching the near-maximum utility (near-maximum utility is a utility close to that of a worker with full information at L26) as the income levels change.

reasonably implemented for the next 2-3 months of work without requiring employers to furnish complete information at the start of their tenure. Notably, for higher-income workers, less future information is needed to approach the utility values of someone with complete information, as observed by the leftward movement of the orange triangles across all lookahead levels.

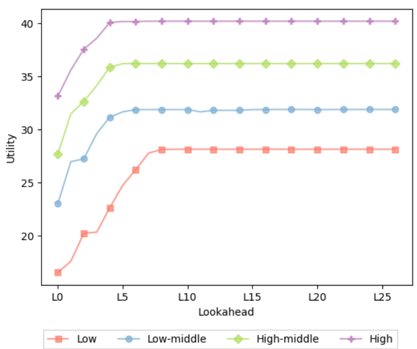
5.4.3 Dynamics of Asset Appreciation and Depreciation

Our simulated space provides a comprehensive platform for delving into the intricacies of parameters within work scheduling. This exploration includes understanding the behavior of individuals with diminishing assets compared to those experiencing favorable returns on their savings.

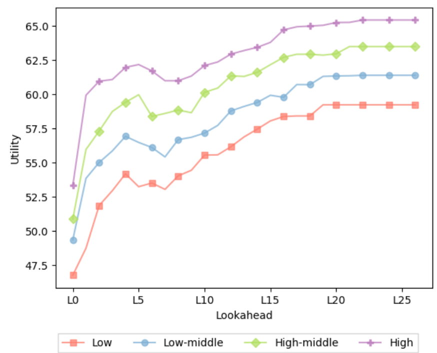
Therefore, in this section, our objective is to investigate the impact of asset appreciation and depreciation, i.e., positive and negative return rates on workers' assets on decision-making across various levels of lookahead. This also serves as an examination of scenarios wherein the workers are already at a (dis)advantage in terms of assets.

Assets depreciate when their value declines over time, influenced by factors like risky investments, fluctuating market conditions, tax obligations, possessions becoming obsolete, and wear and tear, as seen in vehicles, buildings, and cash saved without earning interest. Conversely, asset appreciation occurs when individuals receive returns on their savings or make profitable investments in stocks, among other factors.

All experimental settings are similar to that of §5.4.1 other than the return rates on assets. Here, we compare negative return rates in the range [0.75,0.95] to positive return rates in [1.05,1.25].



(a) Negative return rate analysis



(b) Positive return rate analysis

Figure 5.2: Individuals with similar features but varying levels of lookahead under negative return rates ranging from 0.75 to 0.95 on their assets, as well as positive return rates between 1.05 and 1.25 on their assets.

Analysis. The findings are illustrated in Figure 5.2. Several insights emerge from this analysis.

- Firstly, workers generally experience higher utility when they encounter favorable return rates. This is evident by comparing the utility range in the negative rates plot, which spans from 15 to 40, to the positive plot, which ranges from 45 to 65.
- Secondly, with depreciating assets, individuals benefit significantly from small amounts of lookahead, reaching near-maximum utility. People across income classes achieve a near-maximum utility before Lookahead 10. There is a consistent decline in the lookahead value required to attain near-maximum utility, reaching around lookahead 5 for the highest income class and around 10 for the lowest income class. This observation aligns with the trend observed in the previous section, where more income classes require less future information to reach peak utility values.
- Thirdly, in the case of positive returns, individuals have more flexibility in consumption, as they anticipate overall favorable returns ahead. Future information is not as crucial as in the negative returns scenario, where they are not at a disadvantageous situation with depreciating assets. This explains the late convergence of all income classes to the peak utility value (under positive return rates).

5.4.4 Interventions

In this section, the goal is to examine the mitigating effects of two intervention scenarios. In terms of intervention policies, simulation provides a valuable sandbox environment for testing that might be challenging or even impossible to explore in the real world. In a simulated work scheduling setting, researchers, policymakers, and practitioners can experiment with various interventions, assess their effects, and fine-tune strategies without real-world consequences.

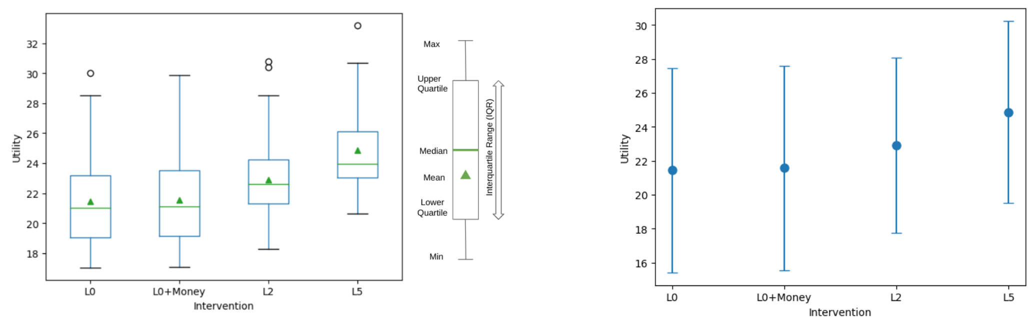


Figure 5.3: The box plot (left) and the error bar plot (right) of the statistical distribution of additional gained utilities for various interventions. The interventions considered are L0 + Money, which involves compensation fees for individuals with no future information; L2, assigning at least two weeks of lookahead to all; and L5, assigning at least five weeks of lookahead to everyone. In the box plot, the green triangle represents the mean, and the green line represents the median. In the error bar plot, the blue circle represents the mean and the error bars are 2 standard deviations from the mean (95% confidence interval).

All experimental settings are similar to §5.4.1. We assign interventions (to a sample of 50 individuals per intervention), as follows:

- **Compensation:** Workers will be compensated for sudden schedule changes and *on-call* shifts. This policy is modeled after the measures approved by the San Francisco Board of Supervisors, which introduced new protections for retail workers in the city, necessitating employers to provide compensation for unpredictable schedules based on factors such as employment type, hourly rate, hours of work [70]. Inspired by this policy, we disburse twice the amount of earnings back when there is a negative shock.

- **Mandatory minimum advance notice:** Every worker should be entitled to a mandated minimum lookahead, meaning they should be aware of their schedule for the upcoming two weeks. This proposition draws inspiration from the Schedules that Work Act of 2014 (H.R. 5159) presented in Congress, which stipulates that if there are alterations to the schedule and minimum hours, the employer must inform the employee at least two weeks prior to the start point of the new schedule [70]. Inspired by this policy, we assign workers two weeks and roughly one month (5 weeks) of minimum lookahead, separately.

Analysis. The results for the intervention scenarios are depicted in Figure 5.3 as a box plot to show the distributions (and a separate corresponding error bar plot to capture uncertainty). Three insights can be derived.

- Firstly, interventions have a positive impact overall, evident from the increased utility across all statistical metrics (mean, median, confidence interval, and all quartiles) post-intervention.
- Secondly, while compensation in the form of additional income for unanticipated schedule changes is beneficial, it does not substitute for providing workers with advance notice of their shift schedules. Having knowledge of future plans with the same income but a predictable schedule appears to be more effective for workers' financial well-being than an unforeseen, volatile schedule compensated with fees. A comparison between *L0 + Money* and *L2* indicates that even incorporating two weeks of lookahead is more advantageous than providing compensation for instability without any lookahead.
- Thirdly, offering just one month of advance notice results in a notable increase in utility, as seen in the comparison between *L5* and *L0*. Even small amounts of advance notice can significantly enhance utility.

Note on the Robustness of Experimental Parameters/Results. It is important to mention that we conducted these experiments with various random seeds, considering different runs (e.g., with a median asset of approximately $70,000 [9], representing the median working-class wealth, as opposed to the population median), and with other realistic discount and return parameters, as well as varied plausible shock sizes. The overall trends in our results in the previous sections remain consistent as long as the chosen parameters fall within more realistic ranges. If one opts for more extreme and unrealistic parameters, the distinctions observed in §5.4.2, 5.4.3, and 5.4.4 will become even more pronounced.

5.5 Chapter Summary

The primary contribution of this chapter lies in the development of an online framework that delves into the financial insecurity stemming from work schedule instability. We provide analytical insights into how lookahead significantly enhances individuals' utility, with a focus on the effectiveness of increased lookahead. Our empirical investigation employs simulations and explores two distinct intervention strategies aimed at mitigating the adverse effects of schedule instability. While our model is being deployed in a work setting, it applies to any scenario that involves temporal uncertainty with the possibility of improvement with lookahead.

In the next chapter, we will completely shift our perspective from individual level to societal level.

Chapter 6

Feedback in Societal-level Systems

Adapted from: Lydia Reader, Pegah Nokhiz, Cathleen Power, Neal Patwari, Suresh Venkatasubramanian, and Sorelle Friedler. "Models for understanding and quantifying feedback in societal systems." In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, pp. 1765-1775. 2022.

Shifting our attention from the individual to the societal level prompts a fundamental question: Why have societal inequities persisted despite extensive activism, educational initiatives, policy changes, and the widespread espousal of values such as equality and non-discrimination?

In our analysis, the persistence of societal inequities can be attributed to the concept of *feedback*. Within every system where inequity endures over time, there exist feedback mechanisms that sustain it. Thus we borrow a tool from systems theory to study this phenomenon on a societal-level *holistically* (as opposed to our agent-based outlooks in studying artificial societies in the previous chapters), as follows.

One of the most extensively studied feedback systems in systems theory is the PID framework [113], consisting of proportional, integral, and derivative forms of feedback. We argue, in this dissertation, with extensive examples and a formal analysis that this simple framework for feedback encapsulates a broad spectrum of societal feedback mechanisms. The framework itself is simple – requiring only a few parameters – and this simplicity is both a value in our ability to interpret what model parameters are telling us about the system, and situates in a favorable place in the tradeoff between model complexity and the need for large amounts of training data. Since society changes over time, it is useful to have a model whose parameters can be accurately estimated with as little history as possible.

An Example: Gender Pay Gaps. The pay gap between men and women is the result of a complex system, impacted by cultural gender roles, biases in the workplace, occupational segregation, and more [4]. While the gender pay gap in the U.S. has reduced over the past 50 years, progress has slowed in recent decades [53]. We note three aspects of the pay gap. First, the current state of inequity is reported on and publicised every year as the pay that an average full-time working woman earns per dollar compared to the average full-time working man.[1] Next, there is a long-term historical inequity, proportional to the sum of this ratio over time, which is the gap in earnings over the lifetime of an average woman retiring today. Finally, there is the short-term change, the change-over-year in the inequity ratio, that represents how fast or slow we are moving towards equity as a society based on this statistic. We refer to these three aspects of the pay gap as the *proportional* (current state), *integral* (historical or long-term), and *derivative* (change). We can identify policies that provide feedback proportional to each term. For example, the salary history question on job applications contributes to keep future salaries (and inequity) close to current salaries and inequity [1], and thus is a proportional feedback mechanism. Gendered career roles are built over time – you are less likely to pursue a career if you do not see many people like you in that career – which is related to the inequity over a lifetime of people who entered that field. We thus see this as an integral mechanism. However, as women enter at higher rates into a profession previously dominated by men, the wages in that profession decrease [117]. Further, reactionary political movements use people's resentments about lost privilege to gain power and reverse policies that helped to reduce the gap [135]. We can see these as derivative mechanisms because the more progress towards equity that is made, the more each effect increases pay inequity.

6.1 Contributions

We can summarize the contributions of this chapter as follows.

- We introduce an approach to model feedback in societal systems using the PID framework (from systems theory).
- We illustrate through numerous examples how the PID framework adeptly represents real-world instances of progress towards, or regressions from, equity.
- We showcase the functionality of this model through three case studies that address historical and persistent inequities.
- We illustrate how the model can be leveraged to assess the impacts of policy changes and interventions.

[1] https://www.aauw.org/resources/article/equal-pay-day-calendar/

6.2 Typology of Feedback in Systems of Inequity

We list 19 specific examples of feedback mechanisms in Table 6.1 which either maintain society's inequity or help to reduce inequity over time. We consider well-publicized inequities in education, employment, criminal justice, political representation, housing, and income. Although many more mechanisms exist, we attempt to give examples that describe a variety of feedback types. In particular, we posit in the rightmost column how next year's inequity is related to current and past inequity:

- *Proportional*: future inequity is a function of current inequity (rows 1-6),
- *Integral*: future inequity is a function of a sum of historical inequity (rows 5-15), or
- *Derivative*: future inequity is a function of the current change in (slope of) inequity (rows 4 and 15-19).

We note that future inequity due to one mechanism can be a function of multiple feedback types. We also note that feedback mechanisms are from different sources, including government policies (rows 5, 7, 11, 18), organization policies (rows 8, 9, 10, 17), laws (rows 2, 14, 19), algorithmic decision system (rows 5, 10, 11, 13), economic rule (rows 7, 12), activism (rows 3, 6), or human psychology (rows 4, 9, 15, 16).

While our analysis of patterns of feedback is systemic and "in the aggregate", we note that the impetus to resist change, or even to adopt particular modes of change that may be ineffective, are often rooted in well-studied group dynamics that typically contribute to the derivative element of feedback. These include a) *backlash*: "The resistance of those in power to attempts to change the status quo is a 'backlash', a reaction by a group declining in a felt sense of power" [131]; b) *reactance*: the pushback when people are confronted with threats to their freedom [23], including not being allowed to discriminate; and, in the case of racial inequity, c) color-blindedness: "Liberalism's very aspirations to color-blindness & equality – while admirable – can impede its goals, as they prohibit race-conscious attempts to right historical wrongs", i.e., change is slowed by them [156].

The purpose of Table 6.1 is to provide many examples of systems of inequity which are maintained and challenged in ways that can be modeled with proportional, integral, and derivative feedback. These mechanisms include existing and potential policies and algorithms, but any change in equity induced by their use would be subject to the other feedback mechanisms of that system. This work provides a simple quantitative model for systemic feedback mechanisms that could be useful in analyzing changes.

Finally, we note what is left out of Table 6.1. There are multiple unpredictable ways in which

Table 6.1: Policies, algorithms, laws, and activism provide feedback in multiple societal systems which exhibit inequities, both to maintain and push back against oppression over time.

#	**System**	**Description**	**Equity?**	**Mechanism**
1	Education	Edu. software is fed unequal student data from oppressive educational contexts; tracking and 'at-risk' labelling keeps students stuck in their current track [129].	Anti	Proportional
2	Employment	The US EEOC 4/5 rule allows legal remedy if, in part, the policy exhibits >20% disparity in hiring within a protected group.	Pro	Proportional
3	Surveillance	Publicity about large inequities in facial recognition by race and gender led to reduced disparities from targeted products [169]	Pro	Proportional
4	Income, Wealth	Support for progressive tax policy change can *decrease* when observing inequity, (e.g., from observing an unhoused person [178]) due to *belief in a just world*.	Anti	Proportional, Derivative
5	Criminal Justice	Since denying parole increases the rate of re-offending after release [203], the current & past racial inequity in parole leads to future inequity in re-offense.	Anti	Proportional, Integral
6	Criminal Justice	The #BlackLivesMatter movement was spurred both by specific incidents of violence and long-term *systemic* violence against Black people [189, 197].	Pro	Proportional, Integral
7	Housing, Wealth	The effects of discriminatory housing policies accumulate over time via lower property value growth in Black and Latinx neighborhoods, which also leads to mortgages with worse terms.	Anti	Integral
8	Higher Ed	Inequity of people admitted to college today will have an effect decades into the future via legacy admits [41].	Anti	Integral
9	Employment	Discrimination in a profession over decades means that there are few examples of a minoritized group in that profession, which then makes members of that group feel less welcome in that profession.	Anti	Integral
10	Employment	Automated hiring models use data from the history of past hires, thus may learn to repeat past discrimination [21].	Anti	Integral
11	Criminal Justice	Future police allocation to an area, and thus future discovered "incidents", is proportional to the cumulative history of incident reports [54].	Anti	Integral
12	Income, Wealth	Excess income (above consumption) adds to wealth in a cumulative sum over time, & earns money on itself (the gross rate of return on wealth) [127].	Anti	Integral
13	Health Care	Algorithms that allocate medical resources to reduce costs assign fewer resources to racial groups who historically received unequal treatment [151].	Anti	Integral
14	Income	The Lilly Ledbetter Fair Pay Act of 2009 allows lawsuits for wage discrimination at one's employer over one's entire career.	Pro	Integral
15	Income, Wealth	Exposure to historical data about *rising* wealth inequality in the U.S. tends to increase support for redistributive policies [132].	Pro	Integral, Derivative
16	Income	As women become a higher percentage of a profession, employers reduce pay to that profession and value it less [117].	Anti	Derivative
17	Higher Ed	The DIF (in SAT future test planning) ensures slow change in race & gender gaps, rather than increasing the use of questions which defy those gaps [179].	Anti	Derivative
18	Voting Rights	Politicians can keep their power despite a changing population by redrawing their district boundaries to include people more likely to vote for them.	Anti	Derivative
19	Voting Rights	Roberts: Voting inequity exists today but is less than it was in the past, thus protection by the Voting Rights Act is not justified [174].	Anti	Derivative

inequity changes over time, e.g., due to a pandemic. In systems theory, this is called the *disturbance* or *process noise*, adding to the feedback as depicted in Figure 1.1, which is distinct from the errors in measuring inequity (which is referred to as *measurement noise*).

6.3 Dynamical PID State Model

We propose a model to quantify proportional, integral, and derivative (PID) feedback mechanisms in systems with inequity. Each PID term is represented as a state variable in the model and is incorporated as feedback on the future state of inequity.

Proportional. Equity is achieved when a societal system produces equal statistics across groups, for example: "racial equity is a state in which race no longer predicts outcomes" [59].

We measure inequity at time n as:

$$x(n) = \frac{\text{outcome measure for people in group A}}{\text{outcome measure for people in group B}} - 1, \tag{6.1}$$

where the "outcome measure" is a societal measure that should be equal across groups if equity is achieved. At equity, the value of $x(n)$ is 0. We choose the group in the numerator so that $x(n)$ is historically above 0, so that readers can consistently interpret $x(n)$ as 'inequity', and work to *reduce* $x(n)$ to 0 as pro-equity. For example, in 1964, from U.S. Census statistics, 70.7% of white Americans (which we set as group A) voted, and 58.5% of Black Americans (group B) voted [200], for a ratio of 1.209 and thus $x(n) = 0.209$. Our choice to subtract 1 in (6.1) is to ensure that minimizing $|x(n)|$ is a desirable goal.

Integral. The cumulative history of inequity is captured by the integral term, $\sigma(n)$. We choose to weigh the most recent history more heavily than the distant history by using an autoregressive filter. Then, the cumulative inequity at time n is given by:

$$\sigma(n) = x(n-1) + \alpha \cdot \sigma(n-1) \tag{6.2}$$
$$\text{with } 0 < \alpha < 1$$

The filter has an infinite impulse response, meaning weights in the cumulative sum will never be completely reduced to zero, but are instead proportional to a factor of α^n at time n. We discuss α further in §6.3.1.

Derivative. The derivative term at time n is the difference in inequity at time n and inequity at time $n-1$:

$$\dot{x}(n) = x(n) - x(n-1) \tag{6.3}$$

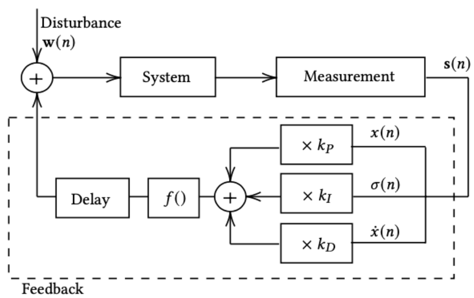


Figure 6.1: A feedback model on a system with inequity $x(n)$, in which the feedback is linear in $x(n)$ itself (the proportional term), in its weighted sum $\sigma(n)$, and in the derivative $\dot{x}(n)$.

State Model We define the state of our model to be the current proportional, integral, and derivative terms at time n, $\mathbf{s}(n) = [x(n), \sigma(n), \dot{x}(n)]^T$. The PID terms are incorporated as feedback in our model, as shown in Figure 6.1.

We describe any other changes to the inequity that is not feedback from the system's outputs as part of the disturbance $\mathbf{w}(n) = [w_0(n), w_1(n), w_2(n)]^T$. We model the dynamics as linear, that is, a weighted sum of the three PID components of the inequity, as well as the disturbance. While it is possible to include non-linearities in the dynamical equations by including an arbitrary function f in the loop as shown in Figure 6.1, in this chapter, we let $f(x) = x$ for simplicity.

Then we model the societal feedback as a linear function of these terms:

$$\mathbf{k}^T\mathbf{s}(n) = k_P x(n) + k_I \sigma(n) + k_D \dot{x}(n), \tag{6.4}$$

where $\mathbf{k} = [k_P, k_I, k_D]^T$, which are the constants which describe how the state evolves. This linear sum, $\mathbf{k}^T\mathbf{s}(n)$ then adds to the current state, specifically, the slope $\dot{x}(n+1)$ at the next time $n+1$ is calculated as the current slope $\dot{x}(n)$ plus this feedback $\mathbf{k}^T\mathbf{s}(n)$ plus some disturbance:

$$\dot{x}(n+1) = \dot{x}(n) + \mathbf{k}^T\mathbf{s}(n) + w_2(n), \tag{6.5}$$

where $w_2(n)$ is the slope disturbance. The system also progresses by: 1) adding the current slope into the inequity for the next time, and 2) keeping track of the cumulative inequity by adding the current inequity to $\sigma(n+1)$. These state update equations are thus:

$$\begin{aligned} x(n+1) &= x(n) + \dot{x}(n) + w_0(n) \\ \sigma(n+1) &= x(n) + \alpha \cdot \sigma(n). \end{aligned} \tag{6.6}$$

These equations (6.5) and (6.6) implement the proportional, integral, and derivative feedback terms as described and justified prior. In short we can write

$$\mathbf{s}(n+1) = A\mathbf{s}(n) + \mathbf{w}(n) \tag{6.7}$$

where

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & \alpha & 0 \\ k_P & k_I & k_D + 1 \end{bmatrix}. \tag{6.8}$$

We note that we do not measure all of the state variables at each time, as we only measure directly the inequity $x(n)$ at each time n. Further, we measure only a noisy version of the inequity $x(n)$. We assume additive noise. To match the typical notation from linear systems, we write $\mathbf{s}(n)$.

$$y(n) = C\mathbf{s}(n) + v(n), \tag{6.9}$$

where $C = [1, 0, 0]$ and $v(n)$ is the measurement noise at time n.

6.3.1 Autoregressive Time Constant Selection

To select a value of the parameter for the autoregressive filter on the integral portion, we choose an appropriate time constant, τ and calculate α as in (6.10).

$$\alpha = e^{-\frac{1}{\tau}} \tag{6.10}$$

The selection of τ is domain-specific and should be based on a reasonable estimate of the time for the impact of historical inequity to decay. For example, if we want terms in the cumulative sum to decay in 10 years and there is one time step per year, $\alpha \approx 0.9$. If we want the terms in the cumulative sum to decay in 100 time steps, then $\alpha \approx 0.99$.

6.3.2 Model Parameter Estimation

Given the dynamic model in (6.7) and a set of longitudinal data for $\{y(n)\}_n$, we want to estimate what parameters of the model are associated with its temporal dynamics. As stated, there is noise in the measurement, and there are disturbances that contribute to the state that are not explained by the PID feedback model. How do we select values for the parameters $\mathbf{k}$ and α from a longitudinal data set?

We provide one method here. Some systems identification methods estimate the entire update matrix A from (6.8), but for our purposes, we only estimate the $\mathbf{k}$ parameters. For our three $\mathbf{k}$ parameters, k_P, k_I, and k_D, we derive a least squares estimator as follows. We define a vector $\Delta\mathbf{s}(n) = \mathbf{s}(n+1) - \mathbf{s}(n)$. An equation for $\Delta\mathbf{s}(n)$ can be written by subtracting $\mathbf{s}(n)$ from both sides of (6.7):

$$\Delta\mathbf{s}(n) = (A - I)\mathbf{s}(n) + \mathbf{w}(n), \tag{6.11}$$

where I is the 3x3 identity matrix. Focusing on the 3rd row of the vector $\Delta\mathbf{s}(n)$, since it is the one element that is a function of the unknown $\mathbf{k}$ parameters,

$$\dot{x}(n+1) - \dot{x}(n) = \mathbf{k}^T\mathbf{s}(n) + w_2(n). \tag{6.12}$$

Defining $\ddot{x}(n) = \dot{x}(n+1) - \dot{x}(n)$, we can then say that:

$$\begin{aligned} \ddot{\mathbf{x}} &= S\mathbf{k} + \mathbf{w}_2, \text{ where,} \qquad (6.13)\\ \ddot{\mathbf{x}} &= [\ddot{x}(1), \ldots, \ddot{x}(N)]^T \\ \mathbf{w}_2 &= [w_2(1), \ldots, w_2(N)]^T \\ S &= [\mathbf{s}(1), \ldots, \mathbf{s}(N)]^T \end{aligned}$$

where we have recorded data from time $n = 0$ to $N + 1$. We could estimate $\mathbf{k}$ in multiple ways, but one easy way would be to use a least-squares approach. Defining superscript $+$ to indicate the pseudoinverse of a matrix,

$$\hat{\mathbf{k}} = (S^T S)^+ S^T \ddot{\mathbf{x}}. \tag{6.14}$$

We note that this estimate is the maximum likelihood estimate in a Gaussian noise case. In short, if we have the full state $\mathbf{s}(n)$ for all times n, we can form the matrix S and vector $\ddot{\mathbf{x}}$ and compute an estimate for $\hat{\mathbf{k}}$.

However, we do not start out with a known state – we only measure $y(n)$ at all times. Thus it is necessary, in order to estimate the parameters $\mathbf{k}$, to first estimate the state $\mathbf{s}(n)$ for all time n. This creates a chicken-and-egg question. A standard approach is to use an expectation maximization (EM) approach to alternately 1) calculate the expected value of the sequence of states $\{\mathbf{s}(n)\}_n$ for all n, and then 2) find the system parameters which maximize the likelihood given the calculated states [66]. In our case, this second part is calculated with (6.14). The first part is described in §6.3.3.

6.3.3 State Estimation

Since we do not measure the state directly or in the absence of noise, our model says that we do not know exactly what the current inequity is, or its slope or cumulative sum. Given a historical set of data measuring the inequity, and known parameters $\mathbf{k}$, we use a Bayesian smoother to estimate the state [183]. We denote this state estimate as $\hat{\mathbf{s}}(n)$ for $n \in \{1, \ldots, N\}$.

As described above, from the state estimates we calculate the change in slope $\ddot{\mathbf{x}}$ which we use with the state S in (6.14) to re-estimate $\mathbf{k}$. We iterate this algorithm until convergence, which we note in practice takes less than 10 iterations.

6.4 Experiments

We test our model on the following real-world datasets:

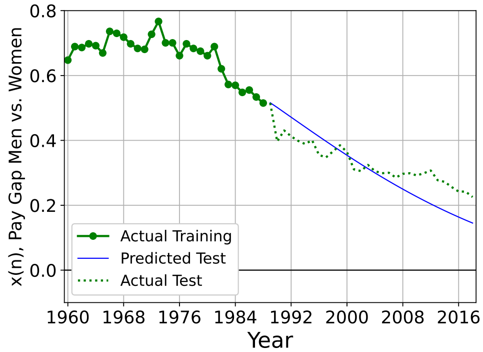
1. *Earnings, Men vs. Women*: The inequity between men and women's earnings is commonly referred to as the gender pay gap, although we note that we do not have a data set inclusive of other genders. For the U.S., we use annual data from 1960 to 2018 [146]. Compiled from U.S. Census data, the data refers to the ratio of median income between men and women full-time, year-round workers. In 2018, women workers' median pay was \$0.82 per dollar of men workers' median pay. Equivalently, we use the inverse, that is, median pay for men divided by the median pay for women, or 1.22, and subtract 1 to obtain an inequity of 0.22.

2. *Voting, White vs. Black*: The percentage of white people who voted divided by the percentage of Black people who voted in the U.S., according to data collected by the U.S. Census Bureau [200]. This data is for national congressional or presidential elections, i.e., every even year, since 1964.

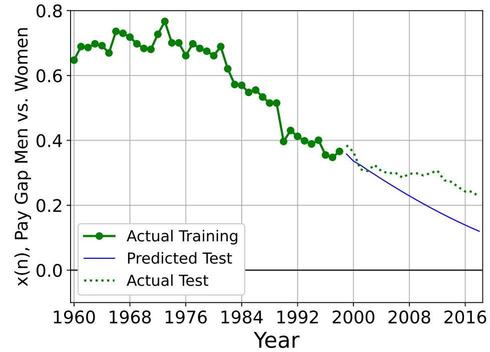
3. *Income, Top 10% of Earners vs. 10% of All Income*: We take the total income of people in the top 10% by income and divide it by 10% of the sum of the income of all people in the U.S. This value is thus a ratio of how much more the people in the top 10% are paid than they would if income was split evenly among all people. The data comes from U.S. tax data collected by Piketty and Saez [176, 162].

We consider the following experimental questions:

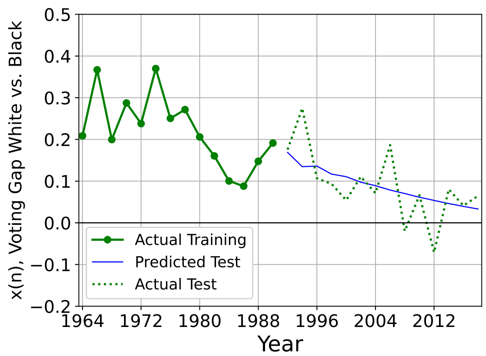
- How well does the model forecast future inequity?
- How can one interpret the model parameters?
- When does the model perform poorly?



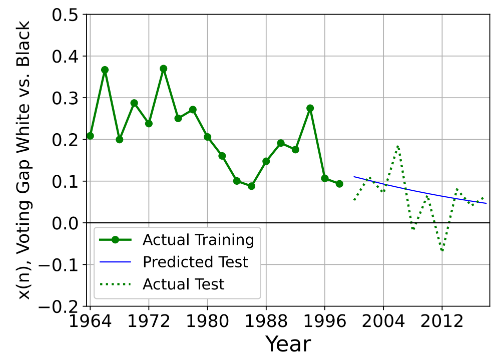
(a1)



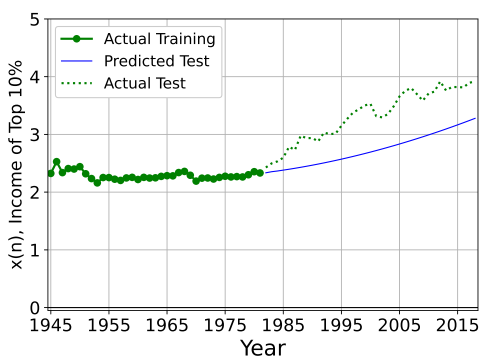
(a2)



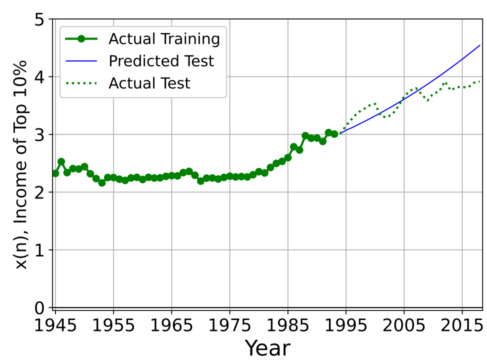
(b1)



(b2)



(c1)



(c2)

Figure 6.2: Training the model only from the historical (Col #1) first 1/2 of data; (Col #2) first 2/3 of the data, we estimate PID model parameters. Then we extrapolate (simulate the model) to forecast the remaining years, and compare to the actual test period data, when $x(n)$ is the U.S. (a) earnings inequity of men vs. women; (b) voting inequity of white vs. Black; (c) income inequity of top 10%. In all plots, $x(n)$ is as defined in (6.1), and a value of 0 (——) is equity.

Data Set #	Name	τ (yrs)	Training Data Used	Proportional k_P	Integral k_I	Derivative k_D
1	Pay Gap Men vs. Women	10	First 1/2	0.0279	-0.0051	-1.07
			First 2/3	0.0372	-0.0066	-1.25
			All	0.0235	-0.0045	-1.08
			Second 1/2	-0.0173	-0.0012	-1.22
2	Voting Gap White vs. Black	20	First 1/2	-0.0511	-0.0075	-1.63
			First 2/3	-0.0729	-0.0033	-1.53
			All	-0.0656	-0.0062	-1.71
			Second 1/2	-0.0595	-0.5278	-0.96
3	Income of Top 10%	100	First 1/2	-0.0150	0.0008	-1.31
			First 2/3	-0.0218	0.0014	-1.25
			All	-0.0129	0.0006	-0.80
			Second 1/2	0.0160	-0.0005	-0.72

Table 6.2: PID model parameters estimated from training. Values are interpreted as: Next year's slope increases by k_P times the current inequity, increases by k_D times the current derivative, and increases by k_I times the current cumulative sum. All parameters with signs that *increase* inequity are red, those that decrease inequity are black, as detailed in §6.4.1.

- How does the model compare to existing simple models?
- How can we use the model to understand possible effects of new policies or algorithms?

6.4.1 Future Inequity Estimation and Parameter Interpretation

In this section, we divide the past into a training period and a test period in order to validate the model's extrapolation performance vs. real world changes in inequity in society. In other words, we essentially pick a threshold year for the purpose of evaluation; the model is trained on the data up to and including the threshold year, and the model then runs, starting with the next year through the present. Since we have data to the present (which was not used in the training) we can see how well the model estimates the "future".

Our results on our three data sets are shown in Figure 6.2. We test (in the left column) using the first half of data for training, and also (in the right column) using the first 2/3 of the data for training. The training is shown as a green solid line, and the actual reserved test data is shown with a green dashed line, and compared to a blue solid line for the model prediction.

We report the estimated model parameters in Table 6.2. We advocate for this model, in part, because the model parameters are interpretable as quantifying feedback types. As detailed in Equations (6.7) and (6.8), next year's slope, $\dot{x}(n+1)$, is k_P times the current inequity $x(n)$, plus k_I times the weighted cumulative sum of inequity, plus $k_D + 1$ times the current slope.[2] In all of our data sets, the current inequity and historical inequity, $x(n)$ and $\sigma(n)$ respectively, are both positive. Thus any $k_P < 0$ or $k_I < 0$ push the model toward forecasting a decrease in inequity in the future, while $k_P > 0$ or $k_I > 0$ would push the model toward forecasting an increase in inequity. However, the sign on the derivative term has a different effect. $k_D < 0$ indicates a push in the model against the current change. If the current slope is negative, the effect of $k_D < 0$ is to forecast slowed down progress towards equity. In contrast, if $k_D > 0$, the derivative feedback reinforces the current direction of change predicted in the system. For each model, we next describe the performance and interpret the parameters.

Earnings, Men vs. Women: The top row of Figure 6.2 shows forecasts for the gender pay gap. We chose an integral term time constant α that corresponds to a decay time of 10 years. The model predicts the downward slope of the data closely when trained on the first half of the data set. When trained on the first 2/3, the model predicts the gender pay gap to decrease more quickly than it actually did. Considering the parameters trained on the entire dataset, the positive sign on k_P indicates a positive association between current inequity and future inequity while the sign of k_I suggests a negative relationship between cumulative inequity and future inequity. The negative sign on k_D is associated with opposition to the current change. In other words, our model finds that only the cumulative income gap is correlated with progress toward equity, while the current pay differences and decrease in the gap over time correlate with feedback against equity.

Voting, White vs. Black: For the voting gap, we chose a time constant such that there is a decay time of approximately 20 years. The voting gap data is particularly noisy, driven in part by different participation rates between presidential election years and non-presidential election years, as well as driven by particular candidates. For example, the 2008 and 2012 elections with President Barack Obama on the ballot had particularly high turnout among Black voters. Nevertheless, the shape of the model forecast closely matches the actual values in the test period. In the voting gap dataset, both k_P and k_I correlate with reduced inequity for every training set used, while the sign of k_D indicates a resistance to the decreasing inequity. Because the actual data oscillates election to election, i.e. a decrease in inequity one timestep is followed by an increase in inequity the next, the magnitude of k_D is generally large. Overall, increasing Black participation in voting (relative to white participation) is found to correlate with resistance.

Income, Top 10% of Earners vs. 10% of All Income: We predicted that the cumulative effect of past inequity would persist much longer into the future in the case of income inequality because excess income is accumulated without loss over time. Therefore, we selected a time constant of

[2]Note the $+1$ comes from the fact that the current slope stays the same in the absence of any feedback.

100 years. Notably, this model does not perform as well when trained on the first half of data, i.e., a period of constant and low inequity from 1945-1981. It may be because the model parameters changed dramatically between the first half and the last half of the data set, coinciding with a wave of tax, benefit, and unionization policy changes starting in the US after 1980 [76]. We can observe, comparing the model parameters when trained on the "First 1/2" vs. on the "Second 1/2", that the proportional parameter switches from negative to positive, while the integral parameter switches from positive to negative. As stated previously, positive values of k_P mean that the slope increases (towards higher inequity) while there is current inequity. We interpret this as saying in the period 1945-1981, the current inequity is associated with a reduction of inequity, while the cumulative history of inequity is associated with increased inequity. In contrast, in the period 1982-2018, our model finds evidence of a dramatic change in the dynamics of income inequality. In this timeframe, the sign of the proportional term indicates a positive relationship with future inequity and the integral term correlates with movement toward a smaller income gap, though with a weaker effect as indicated by the smaller magnitude of k_I. The smaller k_D similarly suggests a weaker relationship between rising income inequality and movement toward equity.

6.4.2 Our Model Compared to Existing Simple Models

In this section, we compare the PID model test forecasts to those generated by other simple regression models that can be learned from a sequence of one-dimensional historical data. We calculate the root mean squared error for linear regression, polynomial interpolation, and decision tree regressors[3] and compare to the PID model. The results for each dataset and training set are shown in Figure 6.3.

Overall, we find that although there is no single model with the lowest error across all datasets and training options, the PID model has one of the lowest error values in general. Given this, as well as the interpretability and manipulability of the model as described previously in this section, we believe that PID is a useful addition to the set of existing simple models.

6.4.3 The Impact of New Policies

The gender pay gap is decreasing, but as discussed previously, the rate of decrease is slowing down. What would the effects be if the nature of feedback in the gender pay gap system could be significantly altered? Consider the example of salary history being used to determine the offered salary for a candidate for a job. The salary history question is believed to perpetuate pay gaps as workers who are currently underpaid tend to get offered lower salaries when taking a new job. For

[3]All are implemented using sklearn's packages and default parameters, with degree of 3 chosen for the polynomial interpolation.

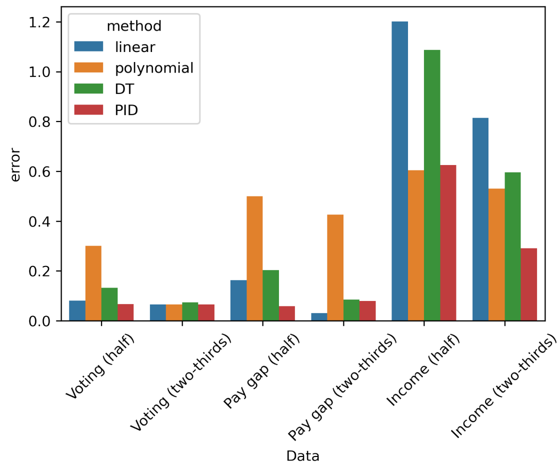


Figure 6.3: RMS errors with respect to test data for each of the comparison methods as well as the PID model for each of the datasets and training combinations.

example, consider that Washington University in St. Louis by policy limits the salary increase for in-university transfers to a maximum of 15%. For people without the privilege to start their career at a highly paid position (often women of color), climbing the ladder can lead to a higher position but, due to salary history policies, sometimes absurdly low pay compared to others in the same job.

When California banned employers from using salary history in 2018, it is estimated that this led to a 1% improvement in the gender pay inequality ratio over a synthetic control in the studied year [78]. We might hypothesize that the feedback effect of the salary history question is proportional – current pay inequity leads directly to future pay inequity. Let us use this hypothesis in the PID framework to investigate the possible long-term effects of the policy under the assumption this policy is the only significant change to the dynamics of the gender pay gap. Let us define the PID model of the pay gap system using the parameter estimates from the entire dataset as shown in Table 6.2, $k_P = 0.0235$, $k_I = -0.0045$, and $k_D = -1.08$. We define the policy to have its own PID terms, k'_P, k'_I, k'_D. If we assume that the salary history ban policy effects only the proportional term, then we can assume $k'_I = 0, k'_D = 0$. To calculate k'_P, we consider the the additive effect of the policy parameters on the system parameters. In 2019, the system with the policy would have a 1% lower output than the system alone. Using (6.6), we derive an equation to solve for k'_P.

$$\frac{x(2018) + \dot{x}'(2018)}{x(2018) + \dot{x}(2018)} = 0.99 \tag{6.15}$$

where $\dot{x}'(n) = \dot{x}(n-1) + (k_P + k'_P)x(n-1) + k_I\sigma(n-1) + k_D\dot{x}(n-1)$. By substituting in the data on the gender pay gap, we find that $k'_P = -0.0536$.

When we simulate the gender pay gap system with and without the salary history ban, we can see that including the policy in the model makes the forecast approach equity much more quickly than the system without the policy. While the policy only results in a 1% decrease in inequity initially, over time, the effects of the policy are expected to be larger. However, it is important to note that the annual 1% improvement in the first year does not continue indefinitely. The PID model here accounts for the existing avenues of societal feedback that, over time, push against the initial change, such that the *slope* in year 2040 is approximately the same with or without the policy.

However, this line of analysis makes some very significant assumptions, namely that this policy alone is the only change in the dynamics of the gender pay gap. It is certainly possible that as one policy is passed that reduces the effects of proportional feedback, the backlash can lead to the creation of new mechanisms to counteract this change, such that effective **k** parameters are not as expected. Progress is not inevitable, and while it certainly good to consider the best-case long-term effects of a policy, it should not be assumed that a single action is enough to dramatically alter a long-lasting system of inequity.

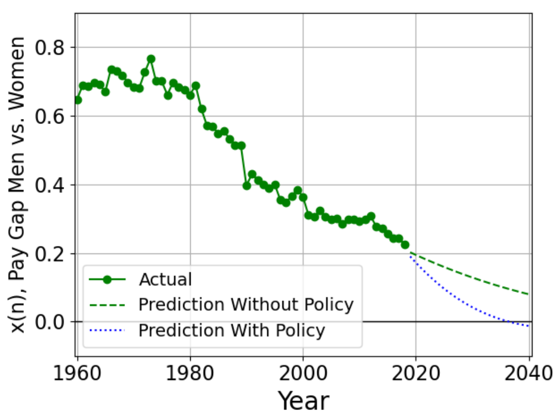


Figure 6.4: Forecasted gender pay inequity with the salary history ban in effect and without the salary history ban.

6.5 Chapter Summary

We present arguments for, and methods to generate, a model for the feedback present in societal systems of inequity. Inequities in outcomes due to racism, sexism, classism, and other systems of oppression are preserved by feedback mechanisms which maintain the status quo, and are reduced by mechanisms which push to address disparities. We build a model with proportional, integral, and derivative feedback terms, and show how historical data can be used to estimate the model's parameters, which then quantify how much of each type of feedback exists in society. We use the model to forecast future trajectories of the inequity and compare our model to alternatives, and show that the error is generally lower than other simple modeling methods. The parameters represent the particular mechanisms, which if changed, would quantitatively alter the trajectory. The model thus introduces a connection between linear systems theory and systemic oppression which could be useful in the modeling and analysis of policy and other mechanisms designed to address social inequity.

Answering the question, "when will we reach equity?" is not just an exercise. U.S. Supreme Court Justice Sandra Day O'Connor, writing the majority opinion in Grutter v. Bollinger that preserved affirmative action, wrote that the "Court expects that 25 years from now, the use of racial preferences will no longer be necessary" [201]. That was 19 years ago; racial inequity in college admissions persists, and O'Connor has since said "That may have been a misjudgement" [198]. But we hope that the model we introduce can do more than provide a likely imperfect window into the future. By explicit reporting of the proportional, integral, and derivative terms, we posit that those seeking equity and those interested in projecting into the future may be better able to reason about the relative impacts of current inequity, longstanding and accumulated inequity, and resistance to – or support for – change.

Chapter 7

Concluding Remarks

In this chapter, we summarize this dissertation's contributions and discuss potential future directions.

7.1 Summary

Overall, we investigated precarity and inequity from individual and societal perspectives. From an individual outlook, we investigated the issue of compounded decisions' effects on individual instability (precarity) by employing a consumption model and Markov Decision Process model to examine how individuals behave in the face of financial shocks. We then expanded this work and studied the long-term effects of financial shocks (in particular, for gig workers with more volatile employment versus office workers with more stable income and employment) with a more realistic model that incorporated realistic constraints such as the desire to avoid ruin, considering expenses for basic needs and time of death.

In both contexts, we explored the efficacy of various policy interventions in mitigating the adverse effects of financial shocks on precarity and showed their effectiveness.

Shifting the focus to the temporal aspect of work schedule instability, we unveiled how inadequate advance notice of schedule changes can profoundly impact workers' financial well-being. By enhancing existing consumption models with an online adaptive algorithm, we explored the effects of future lookahead on workers' financial stability both theoretically and empirically.

Lastly, we transitioned to a societal perspective to study the persisting inequity in complex social systems holistically. We believe there is persisting inequity (despite years of education, activism, and policy changes) because of inequity in feedback mechanisms that are put into place to make it survive. To capture these feedback mechanisms which are too complex to be studied otherwise, we borrow a tool from systems theory that studies feedback mechanisms in complex systems, namely proportional-derivative-integral (PID) framework. PID has three disaggregated feedback terms and states future inequity in terms of current inequity, the sum of historical inequity, and the slope of change in current inequity. This framework helps quantify and study inequity in feedback systems and use mitigation strategies in various cases of historical and current persistent inequity such as income gap and pay gap.

7.2 Future Work

Although our work lays the groundwork for investigating the issues described in §7.1, there are some limitations and potential future research directions.

7.2.1 Group Inequality and Precarity

Thus far, all simulations use parameters that apply to the entire population of households. While individuals in the simulation start at different states, incomes, wealth, and health, the model is identical for every individual. Yet, it is clear that different demographic groups experience bias and discrimination in society. Historical discrimination has a huge impact on an individual's starting point [14, 74], and ongoing discrimination has a major impact on one's income even for the same job (via gender [20] or ability [209]), income shocks [224], health shocks [87], and wealth shocks, as well as the likelihood of benefiting from public policy. In a sense, one's demographic group changes how the economic model must be run. One might apply this model to members of a particular intersectional group who benefit from or are harmed by discrimination in the same manner. Future work could also use multiple models, one each for individuals in different intersectional groups, to show how model differences can lead to disparate results by groups in precarity and financial positions in society.

7.2.2 Heterogeneity and Interconnections among Social Measures

Our framework, like any simulation framework, is limited by what we have chosen to retain and omit. Our simulation assumes societal homogeneity – all individuals are subject to the same forces and constraints, outside of economic differences. This of course, fails to reflect other forms of inequity in society like the ones described above, and is a direction for further exploration. The system could, however if there is heterogeneity in individuals, take that into account in its decisions.

Our choice of income, net worth and health as state variables is well-grounded in the literature on macroeconomic models. However, our modeling of how income and wealth factors influence and are influenced by health relies on a simple formula equating these factors. This interconnection could be explored and improved in future research.

7.2.3 Other Temporal Aspects and Interventions

To improve our online framework in studying altering work schedules' long-term effects, future enhancements could involve the integration of additional factors influencing financial dynamics, which are correlated with temporal uncertainty. This may include incorporating constraints on retirement savings, distinguishing between risky and riskless assets, and accounting for broader socioeconomic changes, such as alterations in workplace organizational structures or shifts in job locations.

Other potential future directions could investigate targeted interventions within specific work settings or environments (e.g., specific interventions designed for the food and retail sectors), aiming to alleviate the impacts of biased decision-making on individuals over extended time horizons.

7.2.4 Feedback Models and Societal Outlooks

There are further future directions in line with our work on studying feedback in social systems with a PID model, as follows.

Portability trap. Are we falling into the "portability trap" [190]? We train our model for each domain / data set, and do not make assumptions about the particular structure of any one system of inequality. However, we are making model assumptions that may not hold in every case — we do not anticipate that a linear feedback model will be sufficient, or that proportional, integral, and derivative terms are best to model the actual mechanisms that keep systemic inequality in place in every type of inequity.

Perception vs. Reality. We use measurements of inequity as the driver for societal feedback mechanisms. However, people's estimates of the level of inequity are inaccurate in the U.S. and U.K., and their estimates are heavily influenced by how much inequality they see locally [81]. If people support policies based on perceived inequity, and perceived inequity is not proportional to measured inequity, this could affect the accuracy of our model. Our model presumes that, society-wide, future changes in inequity are at some level influenced by present, historical, and past changes in measured inequity.

Multidimensionality of Oppression. We model only one inequity measure at a time. In reality, multiple factors contribute to the totality of oppression [63]. For example, the income gap leads to the wealth gap, which then increases the health equity gap. As another example, inequity in the educational system produces future inequity in the criminal justice system [210] via a mechanism called the school-to-prison pipeline. We can imagine extending the model to include multiple measures, and the feedback between different states, although the model complexity would grow.

Overall, this dissertation aims to deepen the understanding of precarity and inequity across individual and societal levels, laying the groundwork for future inquiries into these critical issues.

Appendix A

An Individual-level Framework to Study Precarity: Additional Details

A.1 MDP Detials

In this section, we describe the exact values we choose for the transitions between the MDP states. The transition matrices follow a financial security logic. That is, the more income a household has, the more financially secure they will be, meaning even if a higher income household incurs a financial shock, they have the means (financial and non-financial assets) to cover for their loss. Lower and to a less noticeable degree the middle income households, on the other hand, are less secure, and given a positive decision, they have more chances of staying in the same state due to the previous liabilities. And with a negative decision, they tend to move to inferior states with higher probabilities compared to higher-income households. We set the middle-income classes' chances to a random 50% chance in most cases to give them better chances of improvement. A household can locally reason to go to worse states, better states, or stay in the same states (due to previous debts) as decisions are assigned to them.

The three inferior states given a financial shock due to a negative outcome are:

1. Burning out their savings (encompassed in their net worth),
2. Selling one of their non-financial assets which are health-related (e.g., selling their vehicle or house and falling back on public transportation in the midst of the pandemic, which results in a reduction in their health index). Note that since the net worth encompasses both financial and

non-financial assets, the value of their net worth does not change here. Rather, transiting from a health-related non-financial asset to a financial one will result in a health index reduction by a factor one score (on the scale of the health indexes in the Census Bureau CPS Annual Social and Economic (March) Supplement 2019).

3. Some of the family member opts out of their health insurance to cover for the debts and liabilities, which in turn reduces their health value by a factor of two scores.

The three positive states given a positive decision are:

1. Adding the additional income to their net worth
2. Opting for a better health care plan which improves their health index by a factor of one score.
3. Simply consuming more

The exact values of transition probabilities are as follows: Given a negative transition, the chances of lower-income classes going to a worse state is 29.6% (we have three options for bad states given a negative outcome) while the chances of them staying in the same state and not incurring any financial shocks is 11.11%. The middle-income class is given 50% chance to stay in the same position given a negative (positive) outcome and a 16% chance of going to either of the inferior (better) states. Higher-income classes have 55% chances of staying in the same state given a negative outcome and 15% chances of going to an inferior position for each of the 3 inferior states (similar probabilities for a positive outcome for lower income classes). Given a positive outcome, the higher income class has 11.11% chances of staying in the same state and 29.6% chances of going to each of the better states. After each step of the decision making one of these choices is randomly sampled for the household. The income, net worth, and health indexes get updated accordingly, and the agents continue to interact with the system and environment in an alternating loop.

A.2 Finding Optimal Consumption in IFP

The IFP model [180, 181, 182, 127] uses an endogenous grid method (EGM) to find the optimal consumption path. That is, the EGM necessitates a grid of savings s_i where each saving is the amount of assets with the consumption subtracted. The grid is utilized to interpolate the optimal consumption function. The basis of the grid is on savings because if the assets are not sufficient, the household would consume them all. Else, the savings will be positive (note that the solution which is considered is the origin-based $a_0 = c_0 = 0$). Also, if $s > 0$, then $c < a$. This implies that we can forgo the maximum in 3.3 and solve the following at each s_i:

$$c_i = (u')^{-1} \left\{ \beta \, \mathbb{E}_z \hat{R} (u' \circ \sigma) \, [\hat{R} s_i + \hat{Y}, \, \hat{Z}] \right\}$$

The endogenous asset grid can be computed using $a_i = c_i + s_i$ once we have tuples $\{s_i, c_i\}$. We can get an approximation of the policy $(a, z) \mapsto \sigma(a, z)$ by interpolating $\{a_i, c_i\}$ at each z (note that $z \in \mathsf{Z}$ so it can be paired with a_i).

Model and Implementation Details. In the current model, the exogeneous state process $\{Z_t\}$ is a multi-state process and transition matrix P. We will also assume that $R_t = \exp(a_r \zeta_t + b_r)$ where a_r, b_r are constants and $\{\zeta_t\}$ is i.i.d. standard normal. The labor income itself is defined on the state, percentiles of income.

Using the endogenous grid, and iterating over the interpolations of the optimal consumption functions, until the consumption function converges to a sufficient level, the implementation derives an approximate optimal consumption function. Note that the optimal consumption function gives the consumption value for a given pair of assets and the state. To derive the exact sequence of the consumption from this, we require information on the sequence of states for the specific instance, which we acquire through the defined process following a Markov-like process. The current implementation then uses that information and the approximate optimal consumption function, as well as the basic needs bounds to showcase the consumption path, i.e., sequence of c_ts and a_ts.

Appendix B

An Agent-based Model to Study Precarity with Realistic Constraints: Additional Details

B.1 Additional Proofs and Results on IFP

Given the generic IFP model, we study different settings where we have different constraints on consumption. We assume that there is a basic needs (minimum subsistence consumption) value that provides a lower bound on consumption leading to the inequality constraint in IFP model by setting $b_t \leq c_t \leq a_t$ for all t, where b_t is a known basic needs parameter. Note that the basic needs we discuss are different from the mainstream optimal consumption paths: previous work always assumes that there are enough assets available at all times to cover basic expenditure (thus the individual would never go bankrupt) [193]. But in our setting, we consider a realistic scenario that due to uncertainty, the individual cannot always act ultimately rationally and have enough assets to cover their minimum subsistence, and hence, the amount of assets could drop to a value below the basic needs.

The constraints (e.g., minimum subsistence) we introduce play a vital role in the consumption behavior of an individual. For instance, in the IFP model, an agent that is maximizing utility can avoid ruin assuming they have any realistic constraint such as minimum subsistence. This is formalized in our Lemma 4.

Lemma 4. *Assuming the income* y_t *(*$y_t \geq 0$*) is drawn from a distribution with known mean (denoted by* y*), if we allow the agent to manage their consumption without any restrictive constraints other than* $0 \leq c_t \leq x_t$*, then under the CRRA utility with* $\gamma_c = \frac{1}{2}$ *the agent always prefers consuming with infinite horizon over going to ruin early. If on the other hand we require the minimum subsistence constraint* $b_t \leq c_t$ *together with the IFP constraint* $c_t \leq x_t$*, then there are instances with no feasible solution even if agents might possess sufficient assets* $x_t + y_t$

The proof of Lemma 4 is as follows,

Proof. Assume you are given parameter β and utility $u(c) = 2 \cdot c^{1/2}$ ($u(c) = \frac{c^{1-\gamma_c}}{1-\gamma_c}$ with $\gamma_c = \frac{1}{2}$). Also, assume $r = 1$ for the current model.

Consider the optimal consumption sequence that you get where the agent goes to ruin at time point T (here, asset ruin is assumed to be the point where the available assets for the next iteration reach 0). Let this be $C = \{\widehat{c}_1, \widehat{c}_2, \ldots, \widehat{c}_T\}$. Note that going to ruin at time T is possible only if $y_{T+1} = 0$ (since $\widehat{c}_T \leq a_T$ and $a_{T+1} \geq y_{T+1}$ and therefore if $Y_{T+1} > 0$ then $a_{T+1} > 0$).

Let ϵ be such that $\widehat{c}_T \geq \epsilon \frac{(1+\beta^2)^2}{4\beta^2}$ (Note that $\frac{(1+\beta^2)^2}{4\beta^2} \geq 1$ for all $\beta > 0$ since this translates to $(\beta^2 - 1)^2 \geq 0$). Consider the amended sequence where $C^* = \{\widehat{c}_1, \widehat{c}_2, \ldots, \widehat{c}_T - \epsilon, \epsilon, \ldots\}$. We can see that with this consumption sequence, the agent does not go to asset ruin at point T since we consume less than $\widehat{c}_T$ and we still have ϵ left.

Let $U = \sum_{t=1}^{T-1} \beta^t \cdot 2 \cdot \widehat{c}_t^{1/2}$. Let the total utility from sequence C be $F(C)$ and the total utility from sequence C^* be $F(C^*)$. We can see that $F(C) = U + \beta^T \cdot 2 \cdot \widehat{c}_T^{1/2}$ and $F(C^*) \geq U + \beta^T \cdot 2 \cdot (\widehat{c}_T - \epsilon)^{1/2} + \beta^{T+1} \cdot 2 \cdot (\epsilon)^{1/2}$.

Given our ϵ, we can see that,

$$
\begin{aligned}
&\widehat{c}_T \geq \epsilon \frac{(1+\beta^2)^2}{4\beta^2} \\
&\implies \widehat{c}_T - \epsilon \geq \epsilon \frac{(1-\beta^2)^2}{4\beta^2} \\
&\implies (\widehat{c}_T - \epsilon)^{1/2} \geq \epsilon^{1/2} \frac{(1-\beta^2)}{2\beta} \\
&\implies \widehat{c}_T - \epsilon + 2\beta\epsilon^{1/2}(\widehat{c}_T - \epsilon)^{1/2} + \beta^2\epsilon \geq \widehat{c}_T \\
&\implies (\widehat{c}_T - \epsilon)^{1/2} + \beta \cdot \epsilon^{1/2} \geq \widehat{c}_T^{1/2}
\end{aligned}
$$

which implies that $F(C^*) \geq F(C)$ and therefore, C^* is a better consumption sequence. This contradicts our earlier assumption that C is the optimal consumption sequence. $\square$

While we have only shown this for a specific choice of γ_c, it should be possible to see similar but more complicated arguments that would work for any concave utility function since they fundamentally behave the same way. Now we can see that the agent can always avoid ruin and optimize utility if allowed variability in consumption (without realistic elements).

That is, in the proof of Lemma 4 the agents are allowed to consume almost nothing (the proof of Lemma 4 works if we allow the agent to consume infinitesimally small amounts). But in real-world scenarios this is unrealistic. Therefore, we add a lower bound for basic needs when considering consumption in the following proposition. We can see that this could lead to early ruin with IFP.

Proposition 5. *Assume we have minimum subsistence constraints, $b_t \leq c_t$ where b_t is the minimum subsistence at time t, as well as an upper bound on the consumption introduced by IFP [127], $c_t \leq x_t$ (also assume return on saving, r is 1). Assume the model behaves under the equation, $x_{t+1} = x_t + y_t - c_t$. Under these constraints, there are instances where the individuals would have no feasible solutions with IFP that could account for minimum subsistence even though the individual might be able to account for it by spending the current income, y_t.*

We will argue the validity of this claim, as follows,

Proof. Consider the case where $x_t < b_t$ for some t. We can clearly see that the constraint $b_t \leq c_t \leq x_t$ can no longer be satisfied so there is no valid solution at that point. Since IFP does not allow borrowing, the agent cannot survive with minimum subsistence. Note that this is true even if $(x_t - b_t) + y_t > 0$. This implies that even though we could have accounted for the lack of assets by borrowing from the available income, the constraints on IFP do not allow this course of action and the agent fails to provide for the required minimum subsistence. This implies that the IFP model would not have an admissible solution and therefore IFP fails at this point. This provides us with the desired result. □

In summary, under the IFP model, the agent can only consume from the assets and cannot use income, and it could force the agent into situations where they cannot satisfy the consumption constraints which they might have been able to fulfill if they were allowed to use their income.

B.2 Detailed Argument on the Model from §4.3.2

In this section, we will analyze how we introduce ruin to IFP and how it leads to our model, as well as the technical details involved in introducing the minimum subsistence constraints to the new model in §4.3.2.

For the rest of the section, we will mainly use the following notations. We will use t for time and x_t, y_t, b_t, c_t for the assets, income, minimum subsistence, and consumption at time t respectively. We define τ_0 be the time to ruin, i.e. $\tau_0 = \inf\{t \,|a_t \leq 0\}$. Also, we define τ_d to be the time of death where τ_d comes from an exponential distribution with parameter γ. We let $u(c)$ be the utility (where u is a concave function) achieved by the consumption value c (we will also assume $u(0) = 0$) and β be the discounted factor in the discounted utility model.

Adding Ruin Constraints to IFP

In this section, we will introduce the ruin constraints to IFP and derive a modified model that we will use thereafter. Let c_t, x_t and y_t be the consumption, assets and income at time t. Let r be the return on assets. We will first modify the IFP to include ruin, minimum subsistence and time of death. Introducing time of death, idea of ruin, and minimum subsistence to IFP gives us the following,

$$\begin{aligned} &\max E\left(\sum_{t=1}^{\min(\tau_d,\tau_0)} \beta^t u(c_t)\right) \\ &s.t.\ x_{t+1} = r(x_t - c_t) + y_t \\ &\qquad b_t \leq c_t \leq x_t \end{aligned}$$

Given that the optimization now terminates at the time of ruin, we need to add some constraints that would allow the agent to control their consumption such that they have an opportunity to avoid ruin. To do this, we will use the same constraint used by [16]. Their work introduces a soft constraint $\mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0)$ where $\phi(x_0)$ is a probability parameter that depends on the initial assets. Note that given this ruin constraint, we can remove the upper bound on c_t since the ruin constraint would prevent the individuals from borrowing without bounds (the task which the upper bound is meant to do)

This gives us,

$$\begin{aligned}
&\max E\left(\sum_{t=1}^{\min(\tau_d,\tau_0)} \beta^{t-1} u(c_t)\right)\\
&s.t.\ x_{t+1} = r(x_t - c_t) + y_t\\
&\quad b_t \leq c_t\\
&\quad \mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0)
\end{aligned}$$

and the continuous relaxation of this leads to,

$$\begin{aligned}
&\max E\left(\int_0^{\min(\tau_d,\tau_0)} \beta^{t} u(c_t)dt\right)\\
&s.t.\ dx_t = ((r-1)x_t - rc_t + y_t)dt\\
&\quad b_t \leq c_t\\
&\quad \mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0)
\end{aligned}$$

As we have seen in §4.3.2, we can see that this can be written as,

$$\max E\left(\int_0^{\tau_0} e^{-\gamma t}\beta^t u(c_t)dt + Pe^{-\gamma\tau_0}\beta^{\tau_0}\right) \tag{B.1}$$

$$s.t.\ dx_t = ((r-1)x_t - rc_t + y_t)dt \tag{B.2}$$

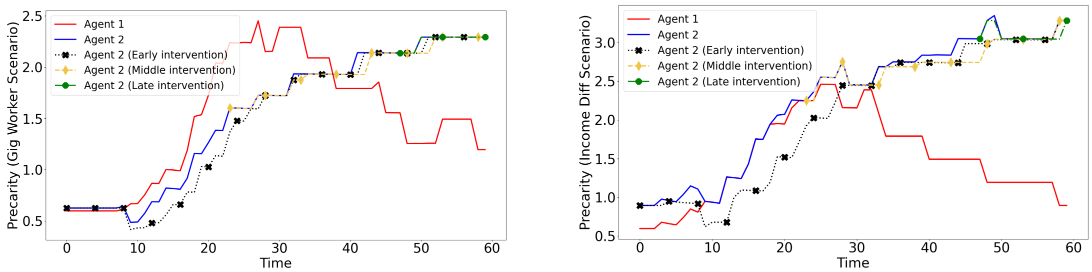
$$b_t \leq c_t \tag{B.3}$$

where P is a Lagrange parameter and $E\left(e^{-\gamma\tau_0}\beta^{\tau_0}\right)$ is the same as $\mathcal{P}[\tau_0 \leq \tau_d] \leq \phi(x_0)$.

Given that we have a proper formulation, the next step we take is to work out a solution for this. We first start off with the equation (B.1), for which we provide a detailed analysis on how to solve and how we can derive an optimal consumption value given the the equation (B.1), in §4.3.2. Next, we provide details on how we can add minimum subsistence using a Lagrange parameter and what this implies for the solution.

Additional Details on §4.3.2 and Adding Minimum Subsistence

In this section, we will mainly provide some additional details on the value function we derived such as how we derive k_1, and then show the modifications we would have when we add minimum subsistence constraints. In §4.3.2, we have shown how we can derive a value function V for the



(a) Gig vs. office worker precarity analysis (b) Minor income difference precarity analysis

Figure B.1: 12-month tax interventions for the same two scenarios studied in §4.5.2. The gig vs. office worker scenario has agents starting with different latent initial instability, similar initial assets, and similar starting income distributions. The colors red and blue correspond to agents 1 and 2, respectively. The minor income difference scenario starts with different initial instability and initial assets ($43,800) with marginally different initial incomes. Agent 1 (red line) has an income of $3,930 and Agent 2 (blue line) has an income of $3,910. The lines with markers in black, yellow, and green represent early, middle, and late intervention start points, respectively.

problem and how we can solve this to get a polynomial that involves $c(x_t)$,

$$x_t = k_1 c(x_t)^{\frac{\gamma_c r}{r-1-\frac{\beta}{2}}} + \frac{\gamma_c r}{\frac{\beta}{2} + (\gamma_c - 1)(r-1)} \cdot c(x_t) - \frac{y_t}{r}$$

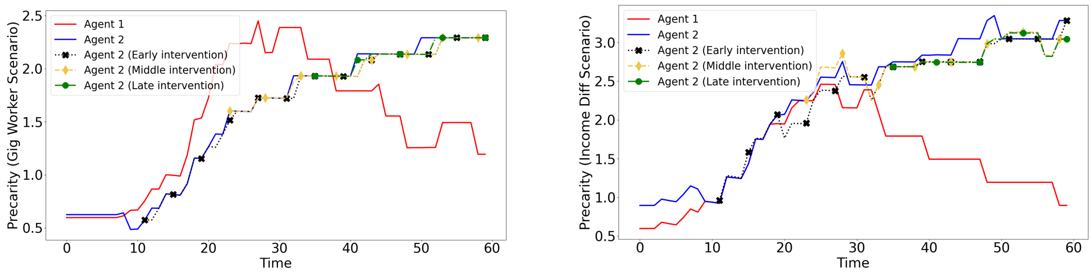
and we can see that the solution to the derived equation satisfies maximization of the value function (since it was derived as a solution to the value function),

$$\beta V(x) = u(c(x)) + \frac{((r-1)x - rc(x) + y)}{r} u'(c(x)) \tag{B.4}$$

For any fixed set of values of γ_c, β, r, finding $c(x_t)$ boils down to solving a polynomial of some specific degree. As we have states before in §4.3.2, given $V(0) = P$, we can also see that,

$$\begin{aligned}\beta P = \beta V(0) &= u(c_0) + \frac{(y_0 - rc_0)}{r} u'(c_0) \\ &= \frac{c_0^{1-\gamma_c}}{1-\gamma_c} + \frac{y_0}{r} c_0^{-\gamma_c} - c_0^{1-\gamma_c} \\ &= \frac{\gamma_c}{1-\gamma_c} c_0^{1-\gamma_c} + \frac{y_0}{r} c_0^{-\gamma_c}\end{aligned}$$

where c_0 is the consumption when $x = 0$ and y_0 is the income at that point (note that since we assume the income process stops at this point in our setting, we can assume y_0 to be 0).



(a) Gig vs. office worker precarity analysis (b) Minor income difference precarity analysis

Figure B.2: Stimulus interventions for the same two scenarios studied in §4.5.2. The gig vs. office worker scenario has agents starting with different latent initial instability, similar initial assets, and similar starting income distributions. The colors red and blue correspond to agents 1 and 2, respectively. The minor income difference scenario starts with different initial instability and initial assets (\$43,800) with marginally different initial incomes. Agent 1 (red line) has an income of \$3,930 and Agent 2 (blue line) has an income of \$3,910. The lines with markers in black, yellow, and green represent early, middle, and late intervention start points, respectively.

Note that, given c_0, and when $\beta/2 \gg r-1$, we get,

$$0 = k_1 + \frac{\gamma_c r}{\frac{\beta}{2} + (\gamma_c - 1)(r-1)} \cdot c_0^{1+\frac{\gamma_c r}{\frac{\beta}{2}-(r-1)}}$$

which gives us the desired k_1.

Minimum Subsistence. In this section, we will try to introduce the minimum subsistence constraint on top of the value function we derived previously. Let x be the assets, $c(x)$ be a function that gives consumption value given x and b be the minimum subsistence value. Note that, given equation B.4 introducing the consumption constraints can be done as follows,

$$\begin{aligned} \max \beta V(x) &= u(c(x)) + \frac{((r-1)x - rc(x) + y)}{r} u'(c(x)) \\ \text{s.t. } c(x) &\geq b \end{aligned}$$

via a modification involving the Lagrange multipliers over the constraints. With this, we get a modified value function $\widehat{V}(x)$ (with Lagrange multiplier λ).

$$\begin{aligned}
\beta\widehat{V}(x) &= \beta V(x) + \lambda(b - c(x)) \\
&= u(c(x)) + \frac{((r-1)x - rc(x) + y)}{r} u'(c(x)) \\
&+ \lambda\,(b - c(x))
\end{aligned}$$

In order to analyze the behavior of this function, we can KKT conditions. With KKT conditions, we get the following. We can see that complementary slackness gives us,

$$\lambda(b - c(x)) = 0$$

Also from the constraint itself, we get $c(x) \geq b$.

Using the stationary conditions we also get,

$$\begin{aligned}
&\frac{\partial\beta\widehat{V}(x)}{\partial x} = 0 \\
&\implies \\
&u'(c(x))c'(x) + \frac{((r-1) - rc'(x))}{r} u'(c(x)) \\
&+ \frac{((r-1)x - rc(x) + y)}{r} u''(c(x))c'(x) - \lambda c'(x) = 0
\end{aligned}$$

which gives us,

$$\lambda = \frac{((r-1)x - rc(x) + y)}{r} u''(c(x)) + \frac{r-1}{r}\frac{u'(c(x))}{c'(x)}$$

Along with complementary slackness, we can see that this implies, $\left[\frac{((r-1)x-rc(x)+y)}{r} u''(c(x)) + \frac{r-1}{r}\frac{u'(c(x))}{c'(x)}\right](b - c(x)) = 0$ which is a well defined differential equation given b. Given this differential equation, we can see that we can still use analytical methods on top of this and find solutions to $c(x)$ given any fixed b. Let c_1 be the solution to $x = k_1 c(x)^{\gamma_c} + \frac{\gamma_c r c(x)}{(\gamma_c - 1)(r-1)} - \frac{y}{r-1}$. In this setting, we can see that we get the consumption,

$$c(x) = \max\{c_1, b\}$$

These functions were calculated using the WolframAlpha[1] symbolic engine. We can see that this gives us an approach to find the $c(x)$ value and therefore a way to define a consumption sequence.

B.3 After-intervention Precarity

The post-intervention precarity plots follow the same pattern as the corresponding asset plots in §4.6. That is, with financial help (tax breaks or stimuli) the agents become less precarious. The more impactful interventions, i.e., earlier and more persistent interventions help the agents become less precarious more prominently than smaller interventions. The results are shown in Figures B.1 and B.2 for both the scenarios studied in §4.5.2 and the corresponding interventions in §4.6.

[1] https://www.wolframalpha.com/

References

[1] Torie Abbott Watkins. The ghost of salary past: Why salary history inquiries perpetuate the gender pay gap and should be ousted as a factor other than sex. *Minn. L. Rev.*, 103:1041, 2018.

[2] Mitsuteru Abe, Fabio Tanaka, Jair Pereira Junior, Anna Bogdanova, Tetsuya Sakurai, and Claus Aranha. Using agent-based simulator to assess interventions against COVID-19 in a small community generated from map data. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '22, page 1–8, Richland, SC, 2022. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450392136.

[3] Rediet Abebe, Jon Kleinberg, and S. Matthew Weinberg. Subsidy allocations in the presence of income shocks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05): 7032–7039, Apr. 2020. doi: 10.1609/aaai.v34i05.6188. URL `https://ojs.aaai.org/index.php/AAAI/article/view/6188`.

[4] Christine Alksnis, Serge Desmarais, and James Curtis. Workforce segregation and the gender wage gap: Is "women's" work valued as highly as "men's"? *Journal of Applied Social Psychology*, 38(6):1416–1441, 2008.

[5] Anne Allison. *Precarious Japan.* Duke University Press, Durham, 2014.

[6] Abhay P Aneja and Carlos F Avenancio-León. No credit for time served? incarceration and credit-driven crime cycles. Technical report, Working paper, 2019.

[7] Jürgen Antony and Torben Klarl. Non-renewable resources, subsistence consumption, and Hartwick's investment rule. *Resource and Energy Economics*, 55:124–142, 2019. ISSN 0928-7655. doi: https://doi.org/10.1016/j.reseneeco.2018.11.002. URL `https://www.sciencedirect.com/science/article/pii/S0928765518300290`.

[8] Soren Asmussen and Hansjorg Albrecher. *Ruin probabilities*, volume 14. World scientific, New Jersey, 2010.

[9] Christian E. Weller Aurelia Glass, David Madland. Unions build wealth for the American working class. *The Center for American Progress*, 2023. URL `https://www.americanprogress.org/article/unions-build-wealth-for-the-american-working-class/#:~:text=A%202022%20survey%20found%20that,in%20that%20period%2C%20and%2011.4`.

[10] James Aylward, Elizabeth Laderman, Luiz E Oliveira, Gladys Teng, et al. How much did the CARES Act help households stay afloat? *FRBSF Economic Letter*, 2021(18):1–06, 2021.

[11] Sinziana Baltatescu. Uncertainty in Neoclassical and Keynesian theoretical approaches: A behavioural perspective. *CES Working Papers*, 7(3):682, 2015.

[12] Abhijit V. Banerjee and Esther Duflo. Inequality and growth: What can the data say? *Journal of Economic Growth*, 8(3):267–299, 2003.

[13] Neil Bania and Laura Leete. Income volatility and food insufficiency in us low-income households, 1991-2003. Technical report, Citeseer, 2006.

[14] Mehrsa Baradaran. *The Color of Money: Black Banks and the Racial Wealth Gap*. Belknap Press of Harvard University Press, 2019.

[15] Nicholas C Barberis. Thirty years of prospect theory in economics: A review and assessment. *Journal of economic perspectives*, 27(1):173–196, 2013.

[16] Erhan Bayraktar and Virginia R. Young. Maximizing utility of consumption subject to a constraint on the probability of lifetime ruin. *arXiv*, 2012. doi: 10.48550/ARXIV.1206.6268. URL `https://arxiv.org/abs/1206.6268`.

[17] Yahav Bechavod, Katrina Ligett, Aaron Roth, Bo Waggoner, and Steven Z Wu. Equal opportunity in online classification with partial feedback. In *Advances in Neural Information Processing Systems*, pages 8972–8982, 2019.

[18] Jess Benhabib and Alberto Bisin. Precarity. *Sharryn Kasmir, The Cambridge Encyclopedia of Anthropology*, 2018. URL `http://doi.org/10.29164/18precarity`.

[19] Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The wealth distribution in bewley economies with capital income risk. *Journal of Economic Theory*, 159:489–515, 2015.

[20] Francine D Blau and Lawrence M Kahn. The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, 55(3):789–865, 2017.

[21] Miranda Bogen and Aaron Rieke. Help wanted: An examination of hiring algorithms, equity, and bias. *Upturn*, 2018.

[22] Jürgen Branke. Artificial societies. *Encyclopedia of Machine Learning, Springer US*, 2010.

[23] Sharon S. Brehm and Jack W. Brehm. *Psychological reactance: A theory of freedom and control*. Academic Press, 2013.

[24] Jenna Burrell. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1):2053951715622512, 2016. doi: 10.1177/2053951715622512. URL `https://doi.org/10.1177/2053951715622512`.

[25] Judith Butler. *Precarious life: The powers of mourning and violence*. Verso Books, London, United Kingdom, 2006.

[26] Judith Butler. *Frames of war: When is life grievable?* Verso Books, London, United Kingdom, 2016.

[27] Monika Bütler. Neoclassical life-cycle consumption: A textbook example. *Economic Theory*, 17(1):209–221, 2001.

[28] Marco Cagetti and Mariacristina De Nardi. Wealth inequality: Data and models. *Macroeconomic Dynamics*, 12(S2):285–313, 2008.

[29] Nancy Cambria, Paul Fehler, Jason Q. Purnell, and Brian Schmidt. Segregation in St. Louis: Dismantling the divide. Technical report, Washington University in St. Louis, 2018.

[30] Christopher F Chabris, David I Laibson, and Jonathon P Schuldt. Intertemporal choice. In *Behavioural and experimental economics*, pages 168–177. Springer, London, 2010.

[31] Yifang Chen, Alex Cuellar, Haipeng Luo, Jignesh Modi, Heramb Nemlekar, and Stefanos Nikolaidis. Fair contextual multi-armed bandits: Theory and experiments. *arXiv preprint arXiv:1912.08055*, 2019.

[32] M. Cristelli, L. Pietronero, and A. Zaccaria. Critical overview of agent-based models for economics, 2011.

[33] Gal Dalal, Assaf Hallak, Steven Dalton, Shie Mannor, Gal Chechik, et al. Improve agents without retraining: Parallel tree search with off-policy correction. *Advances in Neural Information Processing Systems*, 34:5518–5530, 2021.

[34] Greig de Peuter. Creative economy and labor precarity: A contested convergence. *Journal of Communication Inquiry*, 35(4):417–425, 2011. doi: 10.1177/0196859911416362. URL `https://doi.org/10.1177/0196859911416362`.

[35] Angus Deaton. Saving and liquidity constraints. Technical report, National Bureau of Economic Research, 1989.

[36] Angus Deaton. Health, inequality, and economic development. *Journal of economic literature*, 41(1):113–158, 2003.

[37] Angus Deaton et al. *Understanding consumption*. Oxford University Press, Oxford, 1992.

[38] Frederik Michel Dekking, Cornelis Kraaikamp, Hendrik Paul Lopuhaä, and Ludolf Erwin Meester. *A Modern Introduction to Probability and Statistics: Understanding why and how*, volume 488. Springer, London, 2005.

[39] Wouter J Den Haan. Comparison of solutions to the incomplete markets model with aggregate uncertainty. *Journal of Economic Dynamics and Control*, 34(1):4–27, 2010.

[40] Jessica Dickler. With inflation stubbornly high, 58% of Americans are living paycheck to paycheck: CNBC survey, 2023. URL `https://www.cnbc.com/2023/04/11/58percent-of-americans-are-living-paycheck-to-paycheck-cnbc-survey-reveals.html`.

[41] Ezekiel J. Dixon-Román. *Inheriting Possibility: Social Reproduction and Quantification in Education*. University of Minnesota Press, 2017.

[42] Robert Dochow. *Online algorithms for the portfolio selection problem*. Springer, Wiesbaden, Germany, 2016.

[43] Sarah A Donovan, David H Bradley, and Jon O Shimabukuru. What does the gig economy mean for workers?, 2016.

[44] Frederick Douglass. West India Emancipation Speech, Aug. 1857.

[45] Brooke Erin Duffy, Annika Pinch, Shruti Sannon, and Megan Sawey. The nested precarities of creative labor on social media. *Social Media + Society*, 7(2):20563051211021368, 2021. doi: 10.1177/20563051211021368. URL `https://doi.org/10.1177/20563051211021368`.

[46] Cynthia Dwork and Christina Ilvento. Fairness Under Composition. In *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, Wadern/Saarbruecken, Germany, 2018. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH.

[47] Alexander D'Amour, Hansa Srinivasan, James Atwood, Pallavi Baljekar, D. Sculley, and Yoni Halpern. Fairness is not static: Deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*,

FAT* '20, page 525–534, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450369367. doi: 10.1145/3351095.3372878. URL https://doi.org/10.1145/3351095.3372878.

[48] Yonathan Efroni, Gal Dalal, Bruno Scherrer, and Shie Mannor. Beyond the one-step greedy approach in reinforcement learning. In *International Conference on Machine Learning*, pages 1387–1396, Stockholmsmässan, Stockholm SWEDEN, 2018. PMLR, PMLR.

[49] Yonathan Efroni, Mohammad Ghavamzadeh, and Shie Mannor. Online planning with lookahead policies. *Advances in Neural Information Processing Systems*, 33:14024–14033, 2020.

[50] Amir El-Sibaie. Tax Brackets. *Tax Foundation*, 2019. URL https://taxfoundation.org/data/all/federal/2019-tax-brackets.

[51] Hadi Elzayn, Shahin Jabbari, Christopher Jung, Michael Kearns, Seth Neel, Aaron Roth, and Zachary Schutzman. Fair algorithms for learning in allocation problems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 170–179, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450361255. doi: 10.1145/3287560.3287571. URL https://doi.org/10.1145/3287560.3287571.

[52] Vitalii Emelianov, George Arvanitakis, Nicolas Gast, Krishna Gummadi, and Patrick Loiseau. The price of local fairness in multistage selection. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 5836–5842, Macao, 7 2019. International Joint Conferences on Artificial Intelligence Organization. doi: 10.24963/ijcai.2019/809. URL https://doi.org/10.24963/ijcai.2019/809.

[53] Paula England, Andrew Levine, and Emma Mishel. Progress toward gender equality in the United States has slowed or stalled. *Proceedings of the National Academy of Sciences*, 117(13): 6990–6997, 2020.

[54] Danielle Ensign, Sorelle A Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. Runaway feedback loops in predictive policing. In *Conference on Fairness, Accountability and Transparency*, pages 160–171. PMLR, 2018.

[55] Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. "I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in news feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456.

[56] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert,

and Karrie Karahalios. User attitudes towards algorithmic opacity and transparency in online reviewing platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702.

[57] Diana Farrell and Fiona Greig. Paychecks, paydays, and the online platform economy. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 109, pages 1–40, Washington, DC, USA, 2016. JSTOR, National Tax Association.

[58] Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, and J Robert Warren. Integrated public use microdata series, current population survey: version 7.0 [dataset]. *Minneapolis, MN: IPUMS*, 10:D030, 2020.

[59] Forward Through Ferguson Commission. A path to racial equity worksheet, accessed 18 Jan 2022. URL `https://forwardthroughferguson.org/get-involved/pathtoracialequitytool/`.

[60] Peter G. Peterson Foundation. How did Americans spend their stimulus checks and how did it affect the economy?, 2021. URL `https://www.pgpf.org/blog/2021/05/how-did-americans-spend-their-stimulus-checks-and-how-did-it-affect-the-economy`.

[61] Milton Friedman. The permanent income hypothesis. In *A theory of the consumption function*, pages 20–37. Princeton University Press, Princeton, New Jersey, 1957.

[62] Milton Friedman. *Theory of the consumption function*. Princeton university press, Princeton, New Jersey, 2018.

[63] Marilyn Frye. Oppression. In *The Politics of Reality: Essays in Feminist Theory*, pages 1–16. The Crossing Press, 1983.

[64] Peter Ganong and Pascal Noel. Consumer spending during unemployment: Positive and normative implications. *American economic review*, 109(7):2383–2424, 2019.

[65] Lisa A Gennetian and Eldar Shafir. The persistence of poverty in the context of financial instability: A behavioral perspective. *Journal of Policy Analysis and Management*, 34(4): 904–936, 2015.

[66] Zoubin Ghahramani and Geoffrey E. Hinton. Parameter estimation for linear dynamical systems. Technical report, Technical Report CRG-TR-96-2, University of Toronto, Dept. of Computer Science, 1996.

[67] Lesley Gill. *A century of violence in a Red City: Popular struggle, counterinsurgency, and human Rights in Colombia*. Duke University Press, Durham, North Carolina, 2016.

[68] Stephen Gillen, Christopher Jung, Michael Kearns, and Aaron Roth. Online learning with an unknown fairness metric. In *Advances in Neural Information Processing Systems*, pages 2600–2609, 2018.

[69] Marina Gindelsky. Modeling and forecasting income inequality in the United States. *Bureau of Economic Analysis*, 2018.

[70] Lonnie Golden. Irregular work scheduling and its consequences. *Work-Life Balance in the Modern Workplace*, page 115, 2015.

[71] Peter Grandits. An optimal consumption problem in finite time with a constraint on the ruin probability. *Finance and Stochastics*, 19(4):791–847, 2015.

[72] Kathleen Griesbach, Adam Reich, Luke Elliott-Negri, and Ruth Milkman. Algorithmic control in platform food delivery work. *Socius*, 5:2378023119870041, 2019. doi: 10.1177/2378023119870041. URL `https://doi.org/10.1177/2378023119870041`.

[73] Swati Gupta and Vijay Kamble. Individual fairness in hindsight. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 805–806, 2019.

[74] Vivek Gupta, Pegah Nokhiz, Chitradeep Dutta Roy, and Suresh Venkatasubramanian. Equalizing recourse across groups. *arXiv preprint arXiv:1909.03166*, 2019.

[75] Adam M Guren, Alisdair McKay, Emi Nakamura, and Jón Steinsson. Housing wealth effects: The long view. *The Review of Economic Studies*, 88(2):669–707, 2021.

[76] Jacob S. Hacker and Paul Pierson. Winner-take-all politics: Public policy, political organization, and the precipitous rise of top incomes in the United States. *Politics & Society*, 38(2):152–204, 2010.

[77] Anthony Hannagan and Jonathan Morduch. Income gains and month-to-month income volatility: Household evidence from the US financial diaries. *SSRN Electronic Journal*, Online: Paper No. 2659883, 21, 2015. ISSN 1556-5068. doi: 10.2139/ssrn.2659883.

[78] Benjamin Hansen and Drew McNichols. Information and the persistence of the gender wage gap: Early evidence from california's salary history ban. Technical Report 27094, National Bureau of Economic Research, Cambridge, Massachusetts, 2020.

[79] Bradley L Hardy and Trevon D Logan. Racial economic inequality amid the COVID-19 crisis. *The Hamilton Project*, 1:1–10, 2020.

[80] Tatsunori Hashimoto, Megha Srivastava, Hongseok Namkoong, and Percy Liang. Fairness without demographics in repeated loss minimization. In *International Conference on Machine Learning*, pages 1929–1938, Stockholmsmässan, Stockholm, Sweden, 2018. PMLR.

[81] Oliver P. Hauser and Michael I. Norton. (mis) perceptions of inequality. *Current Opinion in Psychology*, 18:21–25, 2017.

[82] Colin B Hawkinson, Sarah B Andrea, Anjum Hajat, Anita Minh, Shanise Owens, Kieran Blaikie, Jessie Seiler, Andrea R Molino, and Vanessa M Oddo. A cross-sectional analysis of work schedule notice and depressive symptoms in the United States. *SSM Popul. Health*, 22 (101413):101413, June 2023.

[83] Ursula Hébert-Johnson, Michael P Kim, Omer Reingold, and Guy N Rothblum. Calibration for the (computationally-identifiable) masses. *arXiv preprint arXiv:1711.08513*, 2017.

[84] Hoda Heidari and Andreas Krause. Preventing disparate treatment in sequential decision making. In *IJCAI*, pages 2248–2254, 2018.

[85] Hoda Heidari, Vedant Nanda, and Krishna Gummadi. On the long-term impact of algorithmic decision policies: Effort unfairness and feature segregation through social learning. In *International Conference on Machine Learning*, pages 2692–2701, Long Beach, California, 2019. PMLR.

[86] Julia R. Henly and Susan J. Lambert. Unpredictable work timing in retail jobs: Implications for employee work–life conflict. *ILR Review*, 67(3):986–1016, 2014. doi: 10.1177/0019793914537458. URL https://doi.org/10.1177/0019793914537458.

[87] Margaret T Hicken, Hedwig Lee, Jeffrey Morenoff, James S House, and David R Williams. Racial/ethnic disparities in hypertension prevalence: reconsidering the role of chronic stress. *American journal of public health*, 104(1):117–123, 2014.

[88] Peter S. Hovmand. *Community Based System Dynamics*. Springer, 2014.

[89] Lily Hu and Yiling Chen. A short-term intervention for long-term fairness in the labor market. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 1389–1398, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee. ISBN 9781450356398. doi: 10.1145/3178876.3186044. URL https://doi.org/10.1145/3178876.3186044.

[90] Lily Hu, Nicole Immorlica, and Jennifer Wortman Vaughan. The disparate effects of strategic manipulation. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 259–268, New York, NY, USA, 2019. Association for Computing Machinery.

ISBN 9781450361255. doi: 10.1145/3287560.3287597. URL https://doi.org/10.1145/3287560.3287597.

[91] Intellingencer. Germany's Kurzarbeit Program, 2020. https://nymag.com/intelligencer/2020/05/germanys-paycheck-protection-program.html.

[92] Shahin Jabbari, Matthew Joseph, Michael Kearns, Jamie Morgenstern, and Aaron Roth. Fairness in reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1617–1626, 2017.

[93] Hyeok Jeong and Robert Townsend. Growth and inequality: Model evaluation based on an estimation-calibration strategy. *Macroeconomic Dynamics*, 12(S2):231, 2008.

[94] Nathan R Johnson and Meredith A Johnson. Precarious Data: Affect, Infrastructure, and Public education. *Rhetoric Society Quarterly*, 50(5):368–382, 2020.

[95] Charles I. Jones. *Consumption*. Stanford Graduate School of Business, Stanford, California, Nov 2020. URL https://web.stanford.edu/~chadj/Consumption2009-11-25.pdf.

[96] Matthew Joseph, Michael Kearns, Jamie H Morgenstern, and Aaron Roth. Fairness in learning: Classic and contextual bandits. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 325–333. Curran Associates, Inc., 2016. URL http://papers.nips.cc/paper/6355-fairness-in-learning-classic-and-contextual-bandits.pdf.

[97] Matthew Joseph, Michael Kearns, Jamie Morgenstern, Seth Neel, and Aaron Roth. Meritocratic fairness for infinite and contextual bandits. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pages 158–163, 2018.

[98] Mark Kac. On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 1949.

[99] Sampath Kannan, Aaron Roth, and Juba Ziani. Downstream Effects of Affirmative Action. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 240–248, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450361255. doi: 10.1145/3287560.3287578. URL https://doi.org/10.1145/3287560.3287578.

[100] Sampath Kannan, Aaron Roth, and Juba Ziani. Downstream Effects of Affirmative Action. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, page 240–248, New York, NY, USA, 2019. Association for Computing Machinery. ISBN

9781450361255. doi: 10.1145/3287560.3287578. URL https://doi.org/10.1145/3287560.3287578.

[101] Jodi Kantor. Working anything but 9 to 5. *The New York Times*, 13, 2014.

[102] Greg Kaplan and Giovanni L Violante. A model of the consumption response to fiscal stimulus payments. *Econometrica*, 82(4):1199–1239, 2014.

[103] Fatih Karahan, Sean Mihaljevich, and Laura Pilossoph. Understanding permanent and temporary income shocks - Liberty Street Economics, Federal Reserve Bank of New York, 2017. URL https://libertystreeteconomics.newyorkfed.org/2017/11/understanding-permanent-and-temporary-income-shocks/.

[104] Ioannis Karatzas, John P. Lehoczky, Suresh P. Sethi, and Steven E. Shreve. Explicit solution of a general consumption/investment problem. *Mathematics of Operations Research*, 11(2): 261–294, 1986. ISSN 0364765X, 15265471. URL http://www.jstor.org/stable/3689808.

[105] Ibram X. Kendi. *How to be an antiracist*. One World, 1 edition, 2019.

[106] Franziska Klügl and Hildegunn Kyvik Nordås. Modelling agent decision making in agent-based simulation – analysis using an economic technology uptake model. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '23, page 1903–1911, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450394321.

[107] Moritz Kuhn. Recursive Equilibria in an Aiyagari-style Economy with Permanent Income Shocks. *International Economic Review*, 54(3):807–835, 2013.

[108] Margot Kushel and Tiana Moore. Towards a New Understanding: The California Statewide Study of People Experiencing Homelessness. *BHHI: UCSF Benioff Homelessness and Housing Initiative*, 2023.

[109] David Laibson, Andrea Repetto, and Jeremy Tobacman. Estimating discount functions with consumption choices over the lifecycle. *National Bureau of Economic Research Cambridge, Mass., USA*, 2007.

[110] Susan J Lambert. Passing the buck: Labor flexibility practices that transfer risk onto hourly workers. *Human relations*, 61(9):1203–1227, 2008.

[111] Susan J. Lambert, Julia R. Henly, and Jaeseung Kim. Precarious work schedules as a source of economic insecurity and institutional distrust. *RSF: The Russell Sage Foundation Journal*

of the Social Sciences, 5(4):218–257, 2019. ISSN 2377-8253. doi: 10.7758/RSF.2019.5.4.08. URL https://www.rsfjournal.org/content/5/4/218.

[112] Alaina Lancaster. The pros and cons of scheduling software. *Restaurant Business*, 2015. URL https://www.restaurantbusinessonline.com/technology/pros-cons-scheduling-software.

[113] Jietae Lee, Su Whan Sung, and In-Beum Lee. *Process Identification and PID Control*. Wiley-IEEE Press, 2009.

[114] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 1603–1612, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456. doi: 10.1145/2702123.2702548. URL https://doi.org/10.1145/2702123.2702548.

[115] Min Kyung Lee, Ishan Nigam, Angie Zhang, Joel Afriyie, Zhizhen Qin, and Sicun Gao. Participatory algorithmic management: Elicitation methods for worker well-being models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, page 715–726, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384735. doi: 10.1145/3461702.3462628. URL https://doi.org/10.1145/3461702.3462628.

[116] Laura Leete and Neil Bania. The effect of income shocks on food insufficiency. *Review of Economics of the Household*, 8:505–526, 2010.

[117] Asaf Levanon, Paula England, and Paul Allison. Occupational feminization and pay: Assessing causal dynamics using 1950–2000 us census data. *Social Forces*, 88(2):865–891, 2009.

[118] Bin Li and Steven CH Hoi. Online portfolio selection: A survey. *ACM Computing Surveys (CSUR)*, 46(3):1–36, 2014.

[119] Bin Li and Steven Chu Hong Hoi. *Online portfolio selection: principles and algorithms*. Crc Press, Boca Raton, 2018.

[120] Fengjiao Li, Jia Liu, and Bo Ji. Combinatorial sleeping bandits with fairness constraints. *IEEE Transactions on Network Science and Engineering*, 2019.

[121] Lan Li, Tina Lassiter, Joohee Oh, and Min Kyung Lee. Algorithmic hiring in practice: Recruiter and hr professional's perspectives on ai use in hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, page 166–176, New York, NY, USA, 2021.

Association for Computing Machinery. ISBN 9781450384735. doi: 10.1145/3461702.3462531. URL `https://doi.org/10.1145/3461702.3462531`.

[122] Lydia T. Liu, Sarah Dean, Esther Rolf, Max Simchowitz, and Moritz Hardt. Delayed Impact of Fair Machine Learning. In *Proceedings of the 35th International Conference on Machine Learning*. PMLR, 10–15 Jul 2018.

[123] Lydia T. Liu, Ashia Wilson, Nika Haghtalab, Adam Tauman Kalai, Christian Borgs, and Jennifer Chayes. The disparate equilibria of algorithmic decision making when individuals invest rationally. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 2020.

[124] Lars Ljungqvist and Thomas J Sargent. *Recursive macroeconomic theory*. MIT press, Massachusetts, 2018.

[125] K Loggins. Here's what happens when an algorithm determines your work schedule. *Vice*, 2020.

[126] Max O. Lorenz. Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219, 1905.

[127] Qingyin Ma, John Stachurski, and Alexis Akira Toda. The income fluctuation problem and the evolution of wealth. *Journal of Economic Theory*, 187:105003, 2020.

[128] C.M. Macal and M.J. North. Tutorial on agent-based modeling and simulation. In *Proceedings of the Winter Simulation Conference, 2005.*, volume 2005, pages 2–15, 01 2005. doi: 10.1109/WSC.2005.1574234.

[129] Michael Madaio, Su Lin Blodgett, Elijah Mayfield, and Ezekiel Dixon-Román. Beyond "Fairness": Structural (In)justice Lenses on AI for education. *arXiv preprint arXiv:2105.08847*, 2021.

[130] Rajib B. Mallick, Michael J. Radzicki, Martins Zaumanis, and Robert Frank. Use of system dynamics for proper conservation and recycling of aggregates for sustainable road construction. *Resources, Conservation and Recycling*, 86:61–73, 2014.

[131] Jane Mansbridge and Shauna L. Shames. Toward a theory of backlash: Dynamic resistance and the central role of power. *Politics & Gender*, 4(4):623–634, 2008.

[132] Leslie McCall, Derek Burk, Marie Laperrière, and Jennifer A. Richeson. Exposure to rising inequality shapes Americans' opportunity beliefs and policy support. *Proceedings of the National Academy of Sciences*, 114(36):9593–9598, 2017.

[133] Bill McCarthy, Angela Carter, Mikael Jansson, Cecilia Benoit, and Ryan Finnigan. Poverty, material hardship, and mental health among workers in three front-line service occupations. *Journal of Poverty*, 22(4):334–354, 2018.

[134] Elaine D McCrate. Unstable and on-call work schedules in the United States and Canada. *International Labour Office, Inclusive Labour Markets, Labour Relations*, 2018.

[135] Heather McGhee. *The sum of us: What racism costs everyone and how we can prosper together.* One World, 2021.

[136] Atif Mian, Ludwig Straub, and Amir Sufi. Indebted demand. *The Quarterly Journal of Economics*, 136(4):2243–2307, 2021.

[137] Jorge Miranda-Pinto, Daniel Murphy, Kieran James Walsh, and Eric R Young. A model of expenditure shocks. *FRB of Cleveland Working Paper*, 2020.

[138] Jorge Miranda-Pinto, Daniel Murphy, Kieran James Walsh, and Eric R Young. Saving constraints, inequality, and the credit market response to fiscal stimulus. *European Economic Review*, 2023.

[139] Thomas M. Moerland, Anna Deichler, Simone Baldi, Joost Broekens, and Catholijn M. Jonker. Think too fast nor too slow: The computational trade-off between planning and reinforcement learning. *arXiv 2005.07404*, 2020.

[140] Dilip Mookherjee and Anthony Shorrocks. A decomposition analysis of the trend in UK income inequality. *The Economic Journal*, 92(368):886–902, 1982.

[141] Jonathan Morduch and Julie Siwicki. In and out of poverty: Episodic poverty and income volatility in the US financial diaries. *Social Service Review*, 91(3):390–421, 2017.

[142] James Morgan. The anatomy of income distribution. *The Review of Economics and Statistics*, pages 270–283, 1962.

[143] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, USA, 2013. ISBN 0511814070.

[144] Hussein Mouzannar, Mesrob Ohannessian, and Nathan Srebro. From fair decision making to social equity. In *Conference on Fairness, Accountability and Transparency*, pages 359–368, 2019.

[145] Hussein Mouzannar, Mesrob I. Ohannessian, and Nathan Srebro. From fair decision making to social equality. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 2019.

[146] National Committee on Pay Equity. The Wage Gap Over Time: In Real Dollars, Women See a Continuing Gap, accessed 2 June 2021. URL https://www.pay-equity.org/info-time.html.

[147] Michael Nau and Matthew Soener. Income precarity and the financial crisis. *Socio-Economic Review*, 17(3):523–544, 06 2017. ISSN 1475-1461. doi: 10.1093/ser/mwx020. URL https://doi.org/10.1093/ser/mwx020.

[148] Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana, Neal Patwari, and Suresh Venkatasubramanian. Precarity: Modeling the Long Term Effects of Compounded Decisions on Individual Instability. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, page 199–208, New York, NY, USA, 2021. Association for Computing Machinery.

[149] Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana, Neal Patwari, and Suresh Venkatasubramanian. Agent-based Simulation of Decision-making under Uncertainty to Study Financial Precarity. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2024.

[150] Pegah Nokhiz, Aravinda Kanchana Ruwanpathirana, Aditya Bhaskara, and Suresh Venkatasubramanian. Counting Hours, Counting Losses: The Toll of Unpredictable Work Schedules on Financial Security, Under Submission.

[151] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[152] Ted O'Donoghue and Jason Somerville. Modeling risk aversion in economics. *Journal of Economic Perspectives*, 32(2):91–114, 2018.

[153] Division of Consumer and Community Affairs. Economic well-being of US households in 2022. *Federal Reserve*, May 2023. URL https://www.federalreserve.gov/publications/files/2022-report-economic-well-being-us-households-202305.pdf.

[154] Board of Governors of the Federal Reserve System. Changes in U.S. Family Finances from 2016 to 2019: Evidence from the Survey of Consumer Finances. *Federal Reserve*, 2019. URL https://www.federalreserve.gov/econres/scf/dataviz/scf/table/#series:Net_Worth;demographic:inccat;population:all;units:median.

[155] U.S. Bureau of Labor Statistics. Consumer Expenditure Surveys. *United States Department of Labor*, 2019. URL https://www.bls.gov/cex/tables.htm.

[156] Ihudiya Finda Ogbonnaya-Ogburu, Angela DR Smith, Alexandra To, and Kentaro Toyama.

Critical race theory for HCI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2020.

[157] George Orwell. *Nineteen eighty-four.* Secker & Warburg, 1949.

[158] Marios Papachristou and Jon Kleinberg. Allocating stimulus checks in times of crisis. In *Proceedings of the ACM Web Conference 2022*, WWW '22, page 16–26, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450390965. doi: 10.1145/3485447.3512047. URL `https://doi.org/10.1145/3485447.3512047`.

[159] Jeffrey Parker. Theories of consumption and saving. *Economics*, 314:1–25, 2010. URL `https://www.reed.edu/economics/parker/s11/314/book/Ch16.pdf`.

[160] Vishakha Patil, Ganesh Ghalme, Vineet Nair, and Y Narahari. Achieving fairness in the stochastic multi-armed bandit problem. *arXiv preprint arXiv:1907.10516*, 2019.

[161] David Pelletier, Simona Bignami-Van Assche, and Anaïs Simard-Gendron. Measuring life course complexity with dynamic sequence analysis. *Social Indicators Research*, 152(3):1127–1151, 2020.

[162] Thomas Piketty and Emmanuel Saez. Income inequality in the United States, 1913–1998. *The Quarterly Journal of Economics*, 118(1):1–41, 2003.

[163] PK. Average, median, top 1%, and all United States household income percentiles in 2020. *DQYDJ*, 2019-2020. URL `https://dqydj.com/average-median-top-net-worth-percentiles/`.

[164] Samuel H Preston. The changing relation between mortality and level of economic development. *Population studies*, 29(2):231–248, 1975.

[165] Jasbir Puar. Precarity Talk: A Virtual Roundtable with Lauren Berlant, Judith Butler, Bojana Cvejić, Isabell Lorey, Jasbir Puar, and Ana Vujanović. *TDR/The Drama Review*, 56 (4):163–177, 2012.

[166] Guillaume Rabault. When do borrowing constraints bind? some new results on the income fluctuation problem. *Journal of Economic Dynamics and Control*, 26(2):217–245, 2002.

[167] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 469–481, 2020.

[168] Hatim A Rahman. The invisible cage: Workers' reactivity to opaque algorithmic evaluations. *Administrative Science Quarterly*, 66(4):945–988, 2021.

[169] Inioluwa Deborah Raji and Joy Buolamwini. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 429–435, 2019.

[170] Lydia Reader, Pegah Nokhiz, Cathleen Power, Neal Patwari, Suresh Venkatasubramanian, and Sorelle Friedler. Models for understanding and quantifying feedback in societal systems. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 1765–1775, 2022.

[171] Michael Reiter. Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665, 2009.

[172] Federal Reserve. Report on the economic well-being of us households in 2015. *Washington, DC*, 2016.

[173] Gilbert Ritschard, Margherita Bussi, and Jacqueline O'Reilly. *An Index of Precarity for Measuring Early Employment Insecurity*, pages 279–295. Springer International Publishing, Cham, 2018. ISBN 978-3-319-95420-2. doi: 10.1007/978-3-319-95420-2_16. URL `https://doi.org/10.1007/978-3-319-95420-2_16`.

[174] C. J. Roberts. Shelby County, Alabama v. Holder, attorney general, et al. *Supreme Court of the United States*, 2013. No. 12-96, Argued February 27, 2013—Decided June 25, 2013.

[175] Alex Rosenblat and Luke Stark. Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International journal of communication*, 10:27, 2016.

[176] Emmanuel Saez. Striking it richer: The evolution of top incomes in the United States (updated with 2017 final estimates). *UC Berkeley*, March 2019. URL `https://eml.berkeley.edu/~saez/TabFig2018prel.xls`.

[177] Paul A Samuelson. A note on measurement of utility. *The review of economic studies*, 4(2): 155–161, 1937.

[178] Melissa L Sands. Exposure to inequality affects support for redistribution. *Proceedings of the National Academy of Sciences*, 114(4):663–668, 2017.

[179] Maria Veronica Santelices and Mark Wilson. Unfair treatment? the case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Educational Review*, 80(1):106–134, 2010.

[180] Thomas Sargent and John Stachurski. Quantitative economics. Technical report, Citeseer, 2014.

[181] Thomas Sargent and John Stachurski. Quantitative economics with Python. *Retrieved from*, 2015.

[182] Thomas J Sargent and John Stachurski. Lectures in Quantitative Economics with Python, 2019.

[183] Simo Särkkä. *Bayesian filtering and smoothing*. Cambridge University Press, 3 edition, 2013.

[184] Jack Schechtman and Vera LS Escudero. Some results on "an income fluctuation problem". *Journal of Economic Theory*, 16(2):151–166, 1977.

[185] Hanspeter Schmidli. On minimizing the ruin probability by investment and reinsurance. *The Annals of Applied Probability*, 12(3):890–907, 2002.

[186] Daniel Schneider and Kristen Harknett. Income volatility in the service sector: Contours, causes and consequences. *The Aspen Institute*, 2017.

[187] Daniel Schneider and Kristen Harknett. Consequences of routine work-schedule instability for worker health and well-being. *American Sociological Review*, 84(1):82–114, 2019.

[188] Daniel Schneider and Kristen Harknett. It's about time: How work schedule instability matters for workers, families, and racial inequality. *The Shift Project Research Brief*, 2019.

[189] Kendra Scott, Debbie S Ma, Melody S Sadler, and Joshua Correll. A social scientific approach toward understanding racial disparities in police shooting: Data from the Department of Justice (1980–2000). *Journal of Social Issues*, 73(4):701–722, 2017.

[190] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 59–68, 2019.

[191] Anuj K Shah, Eldar Shafir, and Sendhil Mullainathan. Scarcity frames value. *Psychological science*, 26(4):402–412, 2015.

[192] Gyoocheol Shim and Yong Hyun Shin. Portfolio selection with subsistence consumption constraints and CARA utility. *Mathematical Problems in Engineering*, 2014:153793, Apr 2014. ISSN 1024-123X. doi: 10.1155/2014/153793. URL `https://doi.org/10.1155/2014/153793`.

[193] Yong Hyun Shin and Byung Hwa Lim. Comparison of optimal portfolios with and without subsistence consumption constraints. *Nonlinear Analysis*, 74(1):50–58, 2011. doi: 10.1016/j.na.2010.08.014.

[194] Harshit Sikchi, Wenxuan Zhou, and David Held. Learning off-policy with online planning. In

Proceedings of the 5th Conference on Robot Learning, pages 1622–1633, London, UK, 2022. PMLR, PMLR.

[195] Guy Standing. The precariat-the new dangerous class. *Amalgam*, 6(6-7):115–119, 2014.

[196] Chad Stone. Fiscal stimulus needed to fight recessions. *Center on Budget and Policy Priorities, April*, 2020. URL https://www.cbpp.org/research/economy/fiscal-stimulus-needed-to-fight-recessions.

[197] Dominique Thomas. Black lives matter as resistance to systemic anti-Black violence. *Journal of Critical Thought and Praxis*, 8(1), 2019.

[198] Evan Thomas. Why Sandra Day O'Connor saved affirmative action. *The Atlantic*, March 2019. URL https://www.theatlantic.com/ideas/archive/2019/03/how-sandra-day-oconnor-saved-affirmative-action/584215/.

[199] Alarith Uhde, Nadine Schlicker, Dieter P. Wallach, and Marc Hassenzahl. Fairness and decision-making in collaborative shift scheduling systems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376656. URL https://doi.org/10.1145/3313831.3376656.

[200] U.S. Census Bureau. Historical reported voting rates, October 2019. URL https://www.census.gov/data/tables/time-series/demo/voting-and-registration/voting-historical-time-series.html.

[201] U.S. Supreme Court. Grutter v. bollinger. *Certiorari to the US Court of Appeals for the Sixth Circuit*, 539(No. 02-241):306, 2003.

[202] Isabel Valera, Adish Singla, and Manuel Gomez Rodriguez. Enhancing the accuracy and fairness of human decision making. In *Advances in Neural Information Processing Systems*, pages 1769–1778, 2018.

[203] Patrice Villettaz, Gwladys Gillieron, and Martin Killias. The effects on re-offending of custodial vs. non-custodial sanctions: An updated systematic review of the state of knowledge. *Campbell Systematic Reviews*, 11(1):1–92, 2015.

[204] Adam Wagstaff and Eddy van Doorslaer. Income inequality and health: What does the literature tell us? *Annual Review of Public Health*, 21(1):543–567, 2000. doi: 10.1146/annurev.publhealth.21.1.543. URL https://doi.org/10.1146/annurev.publhealth.21.1.543. PMID: 10884964.

[205] Peter P. Wakker. Explaining the characteristics of the power (CRRA) utility family. *Health Economics*, 2008.

[206] Philippe Weil. Precautionary Savings and the Permanent Income Hypothesis. *The Review of Economic Studies*, 60(2):367–383, 04 1993. ISSN 0034-6527. doi: 10.2307/2298062. URL `https://doi.org/10.2307/2298062`.

[207] Christian E Weller. Working-class families are getting hit from all sides. *Center for American Progress.*, 2018. URL `https://www.americanprogress.org/article/working-class-families-getting-hit-sides/`.

[208] Min Wen, Osbert Bastani, and Ufuk Topcu. Fairness with dynamics. *arXiv preprint arXiv:1901.08568*, 2019.

[209] Meredith Whittaker, Meryl Alper, Cynthia L Bennett, Sara Hendren, Liz Kaziunas, Mara Mills, Meredith Ringel Morris, Joy Rankin, Emily Rogers, Marcel Salas, et al. Disability, Bias, and AI. *AI Now Institute, November*, 2019.

[210] Maisha T. Winn. *Girl time: Literacy, justice, and the school-to-prison pipeline*. Teachers College Press, 2019.

[211] Alex J Wood. Powerful times: Flexible discipline and schedule gifts at work. *Work, Employment and Society*, 32(6):1061–1077, 2018.

[212] Alex J. Wood. Algorithmic Management: Consequences for Work Organisation and Working Conditions. JRC Working Papers on Labour, Education and Technology 2021-07, Joint Research Centre, May 2021. URL `https://ideas.repec.org/p/ipt/laedte/202107.html`.

[213] Nancy Worth. Making sense of precarity: talking about economic insecurity with millennials in canada. *Journal of Cultural Economy*, 12(5):441–447, 2019. doi: 10.1080/17530350.2018.1485048. URL `https://doi.org/10.1080/17530350.2018.1485048`.

[214] Wenzhi Xi, Zhanfeng Li, Xinyuan Song, and Hanwen Ning. Online portfolio selection with predictive instantaneous risk assessment. *Pattern Recognition*, 144:109872, 2023.

[215] Nutchanon Yongsatianchot, Noah Chicoine, Jacqueline Griffin, Ozlem Ergun, and Stacy Marsella. Agent-based modeling of human decision-makers under uncertain information during supply chain shortages. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2023.

[216] Mark Zandi. The economic impact of a $600 billion fiscal stimulus package. *Brookings Institution*, 2008.

[217] Mark M. Zandi. Assessing the Macro Impact of Fiscal Stimulus 2008. *London: Moody's Economy*, 2008. URL https://www.economy.com/mark-zandi/documents/Stimulus-Impact-2008.pdf.

[218] Angie Zhang, Alexander Boltz, Chun Wei Wang, and Min Kyung Lee. Algorithmic management reimagined for workers and by workers: Centering worker well-being in gig work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. doi: 10.1145/3491102.3501866. URL https://doi.org/10.1145/3491102.3501866.

[219] Xueru Zhang and Mingyan Liu. Fairness in learning-based sequential decision algorithms: A survey. *arXiv preprint arXiv:2001.04861*, 2020.

[220] Xueru Zhang and Mingyan Liu. Fairness in learning-based sequential decision algorithms: A survey. *Handbook of Reinforcement Learning and Control*, page 525, 2021.

[221] Xueru Zhang, Mohammadmahdi Khaliligarekani, Cem Tekin, et al. Group retention when using machine learning in sequential decision making: the interplay between user dynamics and fairness. In *Advances in Neural Information Processing Systems*, pages 15243–15252, Red Hook, NY, USA, 2019. Curran Associates Inc.

[222] Stephan Zheng, Alexander Trott, Sunil Srinivasa, Nikhil Naik, Melvin Gruesbeck, David C Parkes, and Richard Socher. The AI economist: Improving equality and productivity with ai-driven tax policies. *arXiv preprint arXiv:2004.13332*, 2020.

[223] Frederick J. Zimmerman and Michael R. Carter. Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. *Journal of Development Economics*, 2003.

[224] Craig Zwerling and Hilary Silver. Race and job dismissals in a federal bureaucracy. *American Sociological Review*, pages 651–660, 1992.

[225] María J. Álvarez Peláez and Antonia Díaz. Minimum consumption and transitional dynamics in wealth distribution. *Journal of Monetary Economics*, 52(3):633–667, 2005. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2004.06.004. URL https://www.sciencedirect.com/science/article/pii/S0304393205000164.