



Probability, Distributive Justice, and the Promise of Social Insurance

Citation

Friedman, Rachel Zabarkes. 2017. Probability, Distributive Justice, and the Promise of Social Insurance. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:41140203>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Probability, Distributive Justice, and the Promise of Social Insurance

A dissertation presented

by

Rachel Zabarkes Friedman

to

The Department of Government

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Political Science

Harvard University

Cambridge, Massachusetts

March 2017

© 2017 Rachel Zabarkes Friedman

All rights reserved.

Probability, Distributive Justice, and the Promise of Social Insurance

Abstract

Social insurance is the defining policy tool of the modern welfare state, but political theorists have not paid sustained attention to its development or basic character. This dissertation traces the history of social insurance and argues that it is best understood as an evolving distributive regime. Beginning with the origins of mathematical probability theory, I identify a novel account of distributive justice that guided early thinking about life insurance and went on to inspire the first social insurance proposals in 18th-century England and France. I then explain how changes in the understanding of probability encouraged a more collective account of risk and thereby helped support the welfare programs enacted in Western Europe toward the end of the 19th century. Finally, I show how 20th-century social and economic thought undermined the collectivist account of risk from opposing perspectives, and in so doing called the benefits of social insurance into question in ways that have reverberated in both political theory and practice since. Ultimately, my account shows that the prominence and longevity of the insurance principle in modern liberal politics have stemmed from its ability to appeal to competing distributive claims, an appeal that reflects the dualistic character of mathematical probability itself. This fact is both a source of internal tension and a virtue when understood in the proper perspective.

Table of Contents

Introduction	1
I. The problem: justice and chance	1
II. Social insurance as a uniquely modern regime	4
III. A qualified defense of probabilistic social insurance	8
Chapter One: The Origins of Risk and the Growth of Insurance	11
I. The early history of modern insurance	13
A. The emergence of premium insurance	13
B. The theory and practice of early marine insurance.....	17
II. Probability theory and the doctrine of aleatory contracts.....	20
A. The legal background of early probability theory	20
B. Probabilistic justice.....	22
C. Equipossibility and the distributive turn	25
III. Life insurance and probabilistic justice	32
A. Equity in empirical probability theory	32
B. The birth of statistical life insurance	36
C. Mutualism with and without risk.....	41
Chapter Two: Probabilistic Justice and the Beginnings of Social Insurance	45
I. Precursors to social insurance	46
A. Friendly society reform: social insurance writ small	46
B. The first social insurance plans: mutual insurance writ large.....	51
II. The moral character of late-classical probability	58
A. Inverse probability	59
B. The principle of indifference.....	63
C. Mathematical and moral expectation	67
III. Early social insurance in theory and practice	72
A. Between individual choice and social obligation.....	73
B. Causal laws and rational planners.....	79
C. The social insurance moment.....	83
Chapter Three: The Collectivization of Risk and the Early Welfare States	90
I. The rise of the collective view of chance	91
A. Frequentism and its modified case for insurance	91
B. Frequentism and utilitarianism.....	97
C. Frequentist social welfare	105
II. Risk in the early welfare states	108
A. Dilemmas of responsibility	110
B. The targets of social policy	115
C. The flexible actuarialism of early social insurance.....	120
Chapter Four: The Egalitarian Welfare State and the Ambiguities of Insurance	129
I. The egalitarian welfare state emerges.....	131
A. The centrality of insurance to post-war egalitarian welfare	131
B. The amorphous appeal of insurance	136
C. The limits of universal social insurance.....	140
II. Subjective probability and the personalization of chance	144
A. Keynes's transitional account.....	144

B. The rise and triumph of the epistemic view	149
C. The moral character of subjective probability	153
III. The egalitarian welfare state without probability.....	156
A. Probability versus justice	156
B. Rawls's defense of social insurance	159
C. Insurance and distributive theory after Rawls	163
IV. The fate of social insurance in the 20th century and beyond.....	169
Conclusion.....	173
I. The longstanding appeal of insurance	173
II. Explaining the welfare state	175
III. A word for contemporary partisans	179

Introduction

I. The problem: justice and chance

The vulnerability of justice to chance has been the subject of philosophical consideration since the birth of political philosophy. For Plato, the truly just regime depends on the unlikely coincidence of philosophy and political power; for Aristotle, even the best-possible regime requires near-impossible gifts of fortune.¹ The history of modern political thought contains various attempts to formulate a more attainable vision of political justice, one that is among other things less subject to the vicissitudes of chance.

Nevertheless, the challenges posed by uncertain and uncontrollable forces persist on these accounts as well. One powerful modern formulation of the problem starts with the recognition that the outcomes of chancy events frequently result in inequalities that, because they could not have been foreseen or controlled, are inherently unfair to the loser. This formulation and its most influential solution—insurance, and especially social insurance—can be traced directly to mathematical probability theory.²

Probability theory, which first emerged in the mid-17th century, was originally an effort to render contracts with uncertain outcomes fair by equalizing their participants before chance had its way. Such *ex ante* equality, based on each party's mathematical odds of success combined with his potential gain, would ensure that even though their earnings are a product of chance, their agreement is both prudent and morally legitimate. Theorists in this tradition derived their material

¹ See Plato, *Republic*, trans. Robin Waterfield (Oxford: Oxford University Press, 1993), 473d; Aristotle, *Politics*, trans. H. Rackham (Cambridge, MA: Loeb Classical Library, Harvard University Press, 2005), Bk. VII.

² When I refer throughout this text to “probability theory,” I mean both the interpretation of numerical likelihoods and their calculation, which as we will see are inextricably linked.

from games of chance, which thanks to their equally probable outcomes—a fair die, one assumes, is equally likely land on any of its sides—provided the necessarily building blocks for a new science of quantified good sense and distributive justice. Even a wager, as long as mathematically equitable, could on this account be considered both reasonable and fair.

Within the first century of its existence, however, mathematicians working on probability theory began to realize that the dictates of calculation might not always align with those of good sense. A mathematically fair wager could also be a reckless choice for one of the parties, requiring him to pay far more *ex ante* than his likely winning would be worth to him in fact. The concept of utility was born out of the attempt to resolve this problem, and with it emerged a now-centuries-old tradition of trying to reconcile probability's two sides: its account of individual decision-making or reason in the face of uncertainty, and its account of contractual equity or distributive justice among mathematically defined equals.

This same attempt also helped to legitimize life insurance, the first widespread commercial practice to be based on calculations of probability. Insurance was originally a child of marine commerce, and then came to be associated primarily with gambling. Yet beginning in the 18th century, probability theory helped to paint insurance in a new light, as a practice uniquely capable of promoting both individual benefit and the common good. Indeed, group mutual insurance, whether operated for profit or by voluntary associations, appeared to be a model for harmonious social life, in which each member receives his mathematical due and the group as a whole benefits from the security of leaving no one in a state of unjustified dependence or need.

The theoretical basis for this harmony evolved over time, as understandings of probability and the techniques for calculating it changed. The effects of these philosophical developments on the practice of insurance may have been indirect and somewhat limited, but they are nevertheless significant for two reasons. First, as I hope to show, they were central to thinking about the welfare

state, thanks to the idea of social insurance. Understood as mutual insurance writ large, social insurance reflected the original aspirations of modern welfare and proved to be its most important tool. As a result, the changing ways in which insurance was seen to negotiate its obligations to the individual and to the group helped shape the justification for welfare more broadly.

Second, the way in which philosophers and mathematicians understood insurance points up the normative character of probability theory itself. As an attempt to quantify good sense, the calculus rests on (often-implicit) claims about the proper ends of individual choice. As an account of fairness, it rests on (again, often-implicit) claims about the nature of equality and the common good. I submit that such assumptions have motivated thinking about probability throughout its history. They have also shaped important parts of the probability calculus. A central argument of this dissertation is therefore that the quantification of uncertainty via mathematical probability is an inescapably moral and political effort embedded in what appears to be a technical one.

Ian Hacking has famously referred to probability as Janus-faced, in that it has always had both a subjective or epistemic interpretation, which refers to the individual's estimate about a given likelihood, and an objective or aleatory one, which concerns itself with observed facts or physical laws.³ I submit that insurance is correspondingly Janus-faced as well: a reflection both of sound personal judgment and of the aggregate regularity of statistical observations. This feature is, I hope to show, the reason for the enduring political appeal of insurance and a source of its ongoing susceptibility to competing interpretations and political pressures. In short, social insurance is the mirror of probability's unresolved duality.

³ Ian Hacking, *The Emergence of Probability*, 2nd ed. (Cambridge and New York: Cambridge University Press, 2006), 12.

II. Social insurance as a uniquely modern regime

In several key respects, the idea of social insurance can be traced to the birth of modern liberal political thought. Once the purpose of politics is understood as the protection of men's lives and property, and once prudence is defined as the deferral of short-term gratification for the sake of long-term security, it is not a stretch to regard civil society as one great mutual insurer, with payments made in the form of law abidance and taxes and benefits conferred in the form of civil peace and protection. That the political thought of early modern liberalism hinged on the mechanism of contract only heightens the resemblance.⁴ While original accounts of the social contract lack any explicit regard for risk as a calculable entity, in its earliest exponents such as Hobbes and Locke there is nevertheless a demonstrative and even mathematical rigor to the reasoning process that brings about agreement, founded as it is on uncontroversial propositions and the universal good sense that flows from them.⁵

Indeed, insofar as the justice of insurance rests on an abstract but personal possession—the individual's own expectation or risk—which he voluntarily transfers to the insurance society in exchange for the promise of protection, it very much resembles liberal rights theory in its logic. And to the extent that the choice to insure rests as much on the emotional basis of fear as on the counsels of pure mathematics, the similarity is further increased. Yet the first social insurance proposals, made around the time of the French Revolution, also went further than early liberalism in aspiring to replace distributive passions and conflicts with the cool, impartial dictates of the probability calculus. Social insurance is therefore a uniquely modern *distributive* regime. It is

⁴ See also Otto Gierke, *Natural Law and the Theory of Society*, trans. Ernest Barker (Boston: Beacon Press, 1957), 113 fn. 110.

⁵ On several different occasions in his *Essay Concerning Human Understanding*, Locke repeats his view that morality is like mathematics in that both are capable of certain demonstration. See John Locke, *An Essay Concerning Human Understanding*, ed. Roger Woolhouse (London: Penguin, 1997), IV.iv.5; see also II.xxi.3, 8, 42; xxvii.17, 26; xxviii.14; IV.xii.7.

distributive in that it sets out to divide a pool of common resources according to some definition of equality among the relevant parties. It is uniquely modern, I suggest, in that it allows government to harmonize or sublimate competing claims within an apparently technical apparatus defined by the concept of risk.

I develop this argument in four chronologically arranged chapters, each of which covers a period in the history of probability theory. The first chapter catalogues the origins of risk, the birth of mathematical probability, and the early development of insurance. Risk emerged in maritime contracts as a commodity exchanged between two parties, one of whom sought to reduce his exposure to losses and the other of whom sought to profit from the first one's uncertainty. At the time of its first appearance, in roughly the 13th century, risk was quantified but on a strictly prudential basis, as each insurer made his own rough estimates about the likelihood that a particular ship would incur disaster.

Under canon law, such contracts raised the legal and theoretical challenge of how to translate uncertain future profits into fair present prices. Although the rules for doing so were primarily qualitative, they were guided by the well-established principle that each party's gain should be proportional to his investment. Probability theory was born in the mid-17th century out of an attempt to quantify this rule in the context of an interrupted game of chance. Although there has recently been a surge of historical interest in this development, scholars have paid little notice to the fact that early philosophers of probability introduced the language of distributive justice into their mathematical analyses.⁶ They did this by explicitly defining expectation as a party's fair share of a

⁶ A number of fairly recent works have analyzed the legal background of early mathematical probability theory, including Rüdiger Campe, *The Game of Probability: Literature and Calculation from Pascal to Kleist*, trans. Ellwood H. Wiggins, Jr. (Stanford: Stanford University Press, 2012); James Franklin, *The Science of Conjecture* (Baltimore: Johns Hopkins University Press, 2001); Lorraine Daston, *Classical Probability in the Enlightenment* (Princeton: Princeton University Press, 1988); and Ernest Coumet, "La théorie du hasard est-elle née par hasard?," *Annales: Economies, Sociétés, Civilisations* (May-Jun., 1970): 574-598. Although my own interpretation is indebted to these works, particularly Daston's, none of them pays sustained attention to the explicitly distributive focus of many early treatments of mathematical probability.

common pool, or in other words as a kind of distributive entitlement determined by one's potential to win or lose the pot.

This principle of what I call probabilistic justice, according to which a party is entitled to his mathematical expectation as a function of calculated risk, was highly influential in the history of annuities and insurance. Among other things, it helped to legitimate life insurance in England in the mid-1700s, the success of which was in turn seen as proof of the reliability and usefulness of statistical probabilities. Thenceforth, the understanding of insurance as an equitable distributive arrangement would provide a key moral justification for the practice. Yet this same vision coexisted somewhat uneasily with the origins of insurance as a commercial exchange—at times an obviously speculative one, at that—and the more prudential, instrumental logic entailed therein. Indeed, this is the central tension animating the idea of insurance, and a key source of both its political longevity and its apparent contradictions.

Chapter two, which covers the period from roughly the end of the 18th through the mid-19th century, argues that the same account of probabilistic justice that had first legitimated life insurance also inspired both the friendly-society reform movement and the first social insurance proposals in England and France. Mutual and social insurance assume that if enough individuals combine together, paying into the common pool an amount determined by their risks, they can equitably share the burdens of a misfortune that happens to strike any one of them. The fairness of this agreement hinges on the parties' probabilistic equality: Assuming their risks of encountering misfortune are equal, each should pay the same amount to receive the same benefit.

The introduction of statistics into probability calculations introduced a problem, however, namely how to define probabilistic equals in a way that would be both reasonable for the individual and beneficial to the collective. Thinkers of this period, writing in the tradition that I refer to as late-classical probability, addressed the problem using the idea of utility or moral expectation. By touting

the advantages of personal security, they could justify deviations from strict probabilistic justice without undermining the choice-worthy nature of the insurance contract. In retrospect, their accounts would come to appear marred by a number of unsubstantiated assumptions. Yet many of these purported flaws can be explained by the theorists' practical aims, and in particular their desire to preserve the personal relevance of probability values despite the increasingly aggregative or statistical tendencies of the calculus.

By contrast, the alternative account of probability that became influential in the mid-19th century, and which I take up in chapter three, defined mathematical likelihoods as relevant exclusively to statistical groups. This view, known as frequentism, offered a novel account of the relationship between individual judgment and empirical frequencies, according to which the first is completely derivative of the second. The frequentist view had close affinities and historical ties with utilitarianism. It also, I show, accords with the shape and justification of many of the earliest welfare programs instantiated in Western Europe during this time.

Finally, in chapter four, which covers the period from the end of World War II through the late-20th century, I argue that advanced liberal welfare states built on the frequentist idea of social insurance while at the same time attempting, unsuccessfully, to move beyond it. By this time, the more limited risk pooling of the frequentist account no longer seemed adequate to the tasks of welfare, which sought to guarantee self-respect among all citizens regardless of class or status. Yet as long as social policies were still considered "insurance," they continued to appeal to the idea of probabilistic justice and its (at least partial) basis in empirical likelihoods. As a result of this disconnect between the aims of welfare and the workings of insurance, many incisive critics came to think that the insurance rubric was unduly constraining and should be abandoned as a distributive principle.

The account of probability that became dominant during this period also called the virtues of social insurance into question, but for entirely different reasons. The theory of subjective probability, which had direct echoes in mid-century economic thought, differed from nearly every previous account in allowing for the quantification of all types of uncertainty, and not only those for which empirical evidence is available. It did this by understanding all reasoning as a kind of wager, or a bet on what the outcome of an event is likely to be. As a result, subjective accounts effectively collapsed the moral distinction between insurance and gambling, calling earlier theories of social insurance into question at right around the same time that the practice of welfare did the same.

Despite (or perhaps because of) these developments, the most influential work of 20th century political philosophy was a defense of social insurance, or so I argue toward the end of chapter four. Rawls's account of the decision procedure in the original position, and especially the choice of the maximin decision rule, can be read as an effort to preserve insurance as a distributive regime after both social and economic theory had called it into question. Yet while Rawls succeeds in designing the decision procedure in such a way that it is reasonable for participants to insure, I submit that he is less successful in demonstrating why any of us, empirical beings that we are, should support social insurance as a political reality, and with it the liberal welfare state as a whole. I conclude that both Rawls and the developments to which he responded have contributed to political polarization surrounding the welfare state, and suggest that this could be mitigated through a more careful understanding of the historical origins and continued significance of the insurance idea.

III. A qualified defense of probabilistic social insurance

By excluding mathematical likelihoods from the original position, Rawls effectively sidesteps the problem of probability's duality. Because there is no empirical knowledge on which participants can distinguish themselves or identify with a class, there also can be no perspectival dichotomy

between the individual, with his own judgments and aspirations, and the insurer or state, which concerns itself with the common good. In other words, Rawls's design of the veil of ignorance seems to affirm, with reason, that the only incontestable account of probabilistic equality is one that is not probabilistic at all.

In one sense, then, Rawls's social contract is the culmination of the social insurance tradition I recount, and in another a decisive break from that tradition. One of my central claims is that the welfare state emerged as a way reconcile or sublimate probability's two sides—the epistemic or individual and the aleatory or statistical—along with the conflicting distributive claims to which they each give rise. Social insurance promotes a kind of mixed regime that recognizes both personal distinction and equality, individual freedom and the aggregate good. Its determinations are at any given moment contestable, and remain subject to ongoing adjustment in light of partisan claims. Yet it manages to channel such claims rather successfully because, like probability itself, it appeals simultaneously to competing concerns. In imagining insurance without probabilities, I argue, Rawls does away with this dual-faceted appeal, underestimating the psychological sources and political implications of probabilistic reasoning.

In addition, one of the indirect consequences of Rawls's thought has been to encourage a gap between the normative political theory of distributive justice and the empirical political science of the welfare state. While Rawls and many philosophers writing in his wake have intentionally discarded probabilities in the name of an abstract and universal equality, political scientists have emphasized the importance of risk-based solidarity in driving welfare policy.⁷ The latter trend has been fruitful and promises to yield further insights into the electoral politics of welfare. Absent a

⁷ For just a few examples see Torben Iversen, *Capitalism, Democracy and Welfare* (Cambridge: Cambridge University Press, 2005); Philipp Rehm, Jacob S. Hacker, and Mark Schlesinger, "Insecure Alliances: Risk, Inequality, and Support for the Welfare State," *American Political Science Review* 106, no. 2 (2012): 386-406; and Torben Iversen and David Soskice, "Democratic Limits to Redistribution: Inclusionary versus Exclusionary Coalitions in the Knowledge Economy," *World Politics* 67, no. 2 (2015): 185-225.

careful understanding of the moral character of insurance, however, risk-based explanations for welfare will remain disconnected from normative debates about justice. This dissertation therefore aims to help bridge the gap between theoretical and empirical accounts of the welfare state. By considering seriously the nature of probabilistic reasoning and its most characteristic practice, political theorists will be better situated to engage directly with political scientists regarding the promise and limits of social insurance.

Ultimately, my own argument amounts to a qualified defense of probabilistic social insurance. My aim is both to explain its ongoing appeal and to situate it in its proper place as a limited tool for achieving other moral and political aims. Properly understood, social insurance will not appeal to distributive purists, and it will not in itself satisfy those who see politics or social life as aiming at a higher good. It can, however, be a means to such ends. By identifying the competing perspectives and principles that underlie it, and by appreciating the flexible way in which it accommodates them, we will be better situated to understand what a welfare state based on insurance principles can, and cannot, achieve.

Chapter One: The Origins of Risk and the Growth of Insurance

What, exactly, is insurance? In its narrowest sense, insurance is a contract entered into by two parties, one seeking to protect itself against the consequences of a potential loss, the other hoping to benefit from its own relative security by guaranteeing the value of the other's property. Modern insurance involves the exchange of the commodity known as risk, a numerical representation of the possible loss that is transferred from the party facing insecurity to the insurer. In return for shouldering this burden, the insurer receives a premium that depends on both the likelihood and the amount of the loss, which together constitute the risk. In principle, insurers play no role in the actual affairs of those they insure; rather, their involvement is limited to their monetary guarantee and their financial interest is limited to ensuring that the premium they receive is a fair one given the risk.

What we think of today as risk-shifting and risk-spreading arrangements were known long before the advent of modern premium insurance. Greek slave owners could purchase a policy protecting themselves against pecuniary losses resulting from the escape of their slaves.⁸ Roman soldiers had a portion of their pay withheld to provide funds to their families in the event of death in battle, and citizens could join death funds by paying set fees in exchange for a burial allowance. On at least two occasions, the Roman government tried to ensure adequate supplies by indemnifying traders for losses they incurred at sea or due to war. And by the early Middle Ages, cooperative guilds had emerged to provide assistance to members in the event of various contingencies, taken from a common fund to which all members contributed.

All of these proto-insurance arrangements reflect the idea of shifting or spreading potential losses to protect participants from catastrophic outcomes. All specified the relevant dangers and

⁸ These examples come from Karl H. Van D'Elden, "The Development of the Insurance Concept and Insurance Law in the Middle Ages," in *The Medieval Tradition of Natural Law*, ed. Harold J. Johnson (Kalamazoo, Mich: Medieval Institute Publications, Western Michigan University, 1987), 192-95.

promised payment should an unfortunate event come to pass. Yet none shared the most important distinguishing features of modern insurance. First, none explicitly tied the insured party's payment to an *ex ante* quantitative assessment of the danger or loss being insured against.⁹ Instead, they either operated on an *ex post* basis, soliciting contributions or devising compensation after the loss had already occurred, or they linked contributions to compensation but not to the likelihood of the loss.¹⁰ Second, and as a result, none involved the explicit quantification and transfer of the commodity known as risk to a party whose only role is to carry that burden. Although, as we will see, Roman and late-medieval maritime loans did anticipate insurance in imagining the projected loss as a tradable entity, none involved a party whose involvement was limited to shouldering that loss. As a result, such arrangements did not entail the specific aims and challenges of insurance as it later emerged in its modern form.

It may seem needlessly restrictive to define insurance in this way.¹¹ After all, various forms of cooperative scheme have existed for millennia, both before and after the emergence of premium insurance, and were born out of the same impetus that gave rise to the latter. Yet, as I will argue, it was probabilistic life insurance specifically that most directly influenced the future not only of the insurance industry but also of numerous proposals for social welfare on a much broader scale. As a

⁹ For examples, see Charles Farley Trennery, *The Origin and Early History of Insurance* (London: P.S. King & Son, 1926), 107.

¹⁰ In the case of the various efforts by the Roman government to indemnify traders, the merchants paid by simply rendering their services, in exchange for which they received a fixed bounty and a promise that the Emperor would cover losses once they were incurred. Trennery, 119-20. Medieval guilds typically operated by collecting an equal assessment from each member after any one was struck with misfortune. Van D'Elden, 194-95. The Greek slave policy and Roman burial clubs did tie payments out to specific contributions in, but they do not appear to have made any effort to grade those contributions in accordance with the chances of the contingency's occurring.

¹¹ For example, historian Geoffrey Clark includes reversionary annuity societies and redistributive insurance schemes in his consideration of the early life insurance market, despite the fact that unlike premium insurance they did not involve explicit attempts to quantify risks. Yet, as he points out, historically the influence of such schemes was limited, whether due to the inherent limitations of their non-probabilistic methods or to the speculative climate and financial catastrophe that followed quickly in their wake. See Geoffrey Clark, *Betting on Lives* (Manchester and New York: Manchester University Press, 1999), 72.

result, it is fair to say that both the presuppositions and limitations of social insurance as it subsequently developed should be traced to modern statistical insurance, and not to its pre-probabilistic cousins.

I. The early history of modern insurance

A. The emergence of premium insurance

Although there is some dispute among scholars about when exactly maritime insurance emerged in Europe, there is little question that it arose directly out of late-13th-century commercial developments.¹² Prior to this period, merchants shipping their goods by sea had used various partnership arrangements and loans to reduce their financial exposure in the event of a shipwreck.¹³ Early variations on the sea loan conditioned repayment with interest on the safe arrival of the ship, thereby offering the merchant some protection in the event of loss. Such arrangements can be traced back to Roman law, where provision was made for transactions that shifted potential losses between parties for a specified price. Roman maritime loans in particular entailed set rates of interest paid to the lender in exchange for his agreement to forfeit his loan should the voyage fail.¹⁴ Lenders thus acted like insurers in guaranteeing the ship for a price, but because their rates were fixed by statute rather than linked to the particular enterprise, their loans lacked the most significant feature of modern insurance, namely an assessment of the anticipated loss, or risk, to which the guarantee's price is tied.

¹² Florence Edler De Roover, "Early Examples of Marine Insurance," *The Journal of Economic History* 5, no. 2 (1945): 173.

¹³ In addition to the loan-type arrangements discussed here, merchants and ship-owners reduced their exposure to sea risks by owning stakes in a number of different ships, sending their goods on a number of different voyages, and entering into agreements to share one another's losses. See Violet Barbour, "Marine Risks and Insurance in the 17th Century," *Journal of Economic and Business History* 1, no. 4 (1929): 570.

¹⁴ *The Digest of Justinian*, 18.1.35.7, 18.6.1, 19.2.13.5, and 22.2.4, all cited in Franklin, *Science of Conjecture*, 260.

With the revival of Roman law in the 12th century and the subsequent rise of the sedentary merchant, new uncertainty-shifting arrangements began to evolve. In particular, a novel form of sea loan emerged in which ship owners lent money to merchants on the condition that the loan would be repaid on safe arrival of the latter's goods, but forgiven if the goods did not reach their destination. This arrangement not only offered a financial cushion to merchants in the event of a loss, but may also have increased the likelihood of a successful voyage by giving the ship owner a financial interest in protecting the cargo.¹⁵ These loans apparently entailed varying rates of interest that depended on the particular voyage, an anticipation of later risk assessments.¹⁶ Yet they too differed from modern insurance in that they were essentially partnerships between the merchant and the ship owner, each of whom had a direct financial stake in the voyage.¹⁷

The earliest insurance contracts, dating from around the mid-13th century, outwardly resembled sea loans, but in fact modified the underlying arrangement significantly. Rather than loaning the money first and forfeiting it in the event of non-arrival, here the insurer promised to pay a sum of money in the event that a shipment did not arrive safely.¹⁸ In exchange for this promise to shoulder the merchant's loss, the insurer received a premium determined specifically for that voyage and paid by the merchant in advance. The premium, which varied according to the length of the voyage and a number of other factors that might affect its success, was commonly referred to as

¹⁵ De Roover, "Early Examples," 178.

¹⁶ *Ibid.*, 178-79.

¹⁷ Violet Barbour also distinguishes insurance by its involvement of "persons not otherwise interested in the ventures," though she implies that this feature was present before the advent of the premium. Barbour, "Marine Risks," 571. According to De Roover's chronology, however, the direct precursor to premium insurance, the insurance loan, did not involve a third party but rather a merchant and a ship owner each of whom was directly involved in the venture. De Roover, "Early Examples," 185.

¹⁸ De Roover, "Early Examples," 185.

rischio, the linguistic predecessor of risk.¹⁹ Insurance thus differed from its predecessors in two key respects: first, it involved the explicit exchange of an uncertainty, independent of the underlying property being insured; and second, it involved a party, the insurer, whose role in the transaction was limited to his assumption of that risk.

The first of these two distinguishing features is reflected in the introduction of the premium. The premium rested on the assumption that the likely downside of a business venture could be alienated from the underlying property and, its value having been assessed, sold as a commodity. The idea that a probable loss could be alienated and transferred was not the sudden innovation of maritime insurance.²⁰ Yet the insurance premium is the most direct progenitor of the concept of risk as we know it: namely, a measurement of a potential loss made before the relevant contingency occurs and exchanged independently of the good to which it refers. Not only do we owe the term risk directly to maritime insurance's *rischio*, but maritime insurance also appears to have been the first widespread practice to involve the explicit exchange of this commodity.²¹

¹⁹ For example, the 14th century account book of a Florentine firm refers to an 8.75 percent *rischio* charged by a shipping company who had agreed to deliver goods at its own peril but at the buyer's expense. *Ibid.*, 181-82.

²⁰ Both Roman and late medieval maritime loans operated with the understanding, implicit or explicit, that a lender could assume the possible losses of a venture in exchange for compensation in the form of interest. Looking as far back as the *Digest of Justinian*, one finds an acknowledgment that "it is possible for a thing to be deposited, or a mandate to be undertaken, on the condition that the thing is at the risk of him who undertook the deposit or the mandate." In the case of maritime loans in particular, the *Digest* allows that "in exchange for bearing a risk, a person may recover what he has given and something else besides money." *Digest of Justinian*, 17.1.39; 22.2.5, quoted in Franklin, *Science of Conjecture*, 260. Franklin concludes from such evidence the Romans understood risk as "almost a detachable entity in itself," and notes their conceptual advance in seeing "hopes and perils as quantities that can have a price." There are also examples outside the maritime context of early jurists and scholars contemplating the possibility of trading an uncertainty. Among them is a late-13th century legal treatise that, though employing the term *periculum* rather than *resicum* or *rischio*, contemplates selling "the value of the probability or the probable expectation of gain that can be obtained through commercial operations" enabled by a capital investment. G. Todeschini, *Un trattato di economia politica francescana: il De emptioibus et venditionibus, de usuris, de restitutionibus di Pietro di Giovanni Olivi* (Roma: 1980), 110, quoted in Silvain Piron, "L'apparition du *resicum* en Méditerranée Occidentale, XIIe-XIIIe siècles," in *Pour une histoire culturelle du risque*, ed. E. Collas-Heddeland et al. (Strasbourg: Editions Histoire et Anthropologie, 2004), 20.

²¹ I do not have direct evidence for this claim, but I believe it is significant that while the possibility of selling a risk was explicitly contemplated in the late-13th century, legal doctrine only evolved to allow the practice *after* the introduction of insurance. According to historian John Noonan, early usury theory rested on the view that risk was a mark of ownership, not a claim to further profit. A capitalist's earnings therefore derived from the use of his money rather than his acceptance of risk. Yet beginning in the early 15th century—that is, shortly after the birth of insurance—nearly all commentators began to disregard this doctrine and allow that risk could be exchanged independent of title. In other

The understanding of risk as an independently tradable good is closely related to the second feature distinguishing maritime insurance from its predecessors: the introduction of a party not directly involved in the primary venture. While sea loans typically involved partners to the enterprise, insurance agreements were made between merchants and underwriters who had no connection to the business beyond their assumption of the risk.²² This arrangement was made possible by the premium, and in particular the assumptions it reflected about the independence and quantifiability of risk.

The invention of risk as a commodity may have been, at least in part, a response to canon law prohibitions on usury.²³ Maritime loans were clearly problematic because of their use of interest as compensation for the lender.²⁴ Insurance, however, bypassed the loan in favor of an outright sale or *quid pro quo*.²⁵ Although canon law had long regarded the bearing of risk as a marker of ownership and denied that one could profit simply from its assumption, by the late 15th century most commentators were willing to simply overlook this doctrine and accept the exchange of risk

words, it may be the case that legal doctrine evolved to meet the needs of business practice, and had little reason to do so prior to the rise of insurance. See John T. Noonan Jr., *The Scholastic Analysis of Usury* (Cambridge, MA: Harvard University Press, 1957), ch. 10.

²² De Roover notes that once insurance was established as a business, it typically involved multiple independent underwriters insuring each venture. De Roover, "Early Examples," 180, 187-88. For evidence that alternative forms of risk-management in maritime trade involved parties to the business agreement who simply subdivided the potential dangers among themselves, see A. D. M. Forte, "Marine Insurance and Risk Distribution in Scotland before 1800," *Law and History Review* 5, no. 2 (1987): 406.

²³ Franklin, *The Science of Conjecture*, 274.

²⁴ As a result, many loan agreements simply avoided mentioning the interest payment, although scholars assume that some form of compensation was given to the insurer before the voyage. See Humbert O. Nelli, "The Earliest Insurance Contract: A New Discovery," *The Journal of Risk and Insurance* 29, No. 2 (1972): 215-220. For a survey of scholastic treatments of usury in sea loans, see Noonan, *Scholastic Analysis*, 145-151.

²⁵ This is the argument made by Portuguese jurist Santerna to distinguish insurance from maritime loans in his seminal treatise on insurance, *Tractatus de assecurationibus et sponsionibus: nunc primum luce donatus* (Lisbon: Grémio dos Seguradores, 1971), 203-12. Nor was Santerna alone: Beginning at the turn of the 15th century, several leading moralists declared insurance licit because in the sale of risk a burden was assumed and compensation given for its assumption. Noonan, *Scholastic Analysis*, 203.

independent of title.²⁶ As a result, insurance encountered little or no problem with canon law and was readily accepted by 15th- and 16th-century jurists as being in principle a fair exchange.²⁷

B. The theory and practice of early marine insurance

Analytically and legally, the main challenge for insurance was how to fairly price the risk.²⁸ Canon law norms of contractual equity required that parties' financial exposure be proportional to their total share in the profits.²⁹ Yet as in all so-called aleatory contracts, the outcome of which hinges by definition on a contingency, the actual profits to each party from an insurance arrangement were highly uncertain. As a result, the price had to incorporate not only the final amount at stake, but also the likelihood that those profits would be obtained.³⁰ As the Portuguese jurist Santerna explained in his 1488 *On Insurance and Merchants' Bets*, "the insurer sells only the hope of a future outcome... From the fact that this hope is uncertain, it might not seem capable of estimation." Yet it is possible to estimate the just price, not "by how much the thing or goods would be worth in case the peril was realized, but at how much the doubtful event should likely be estimated. In which case the price seems to be constituted with respect to that hope."³¹ In short,

²⁶ One obscure 15th-century scholar objected to insurance on the grounds that one could not transfer a risk without transferring ownership, but his view was quickly drowned out by the chorus of jurists who accepted that the assumption of risk in itself entitled the capitalist or insurer to a reward. Noonan, *Scholastic Analysis*, 203-205.

²⁷ Franklin, *Science of Conjecture*, 274-78.

²⁸ Ibid. and Daston, *Classical Probability*, 210-23.

²⁹ Franklin, *Science of Conjecture*, 262-269 and Noonan, *Scholastic Analysis*, 281-283, 289.

³⁰ Noonan, *Scholastic Analysis*, 147. The defenders of life annuities thus argued that such contracts were equitable as long as the potential for loss or gain was equalized among the parties. For two 14th century examples of this argument, see Franklin, *Science of Conjecture*, 270-72.

³¹ Santerna, *Tractatus*, pt. 5, sections 3 and 4, quoted in Franklin, *Science of Conjecture*, 277.

one must consider both the value of the goods and the likelihood of their arrival (aptly designated the “hope”), and fix the insurance premium accordingly.

In practice, early insurance premiums were based more on guesswork and negotiation than rigorous calculation. The early underwriters were not specialists but merchants, ship owners, bankers, and others who underwrote as a sideline to their principal businesses.³² Moreover, early insurance agreements often looked more like pure wagers than the equitable partnerships of legal treatises.³³ In fact, betting played a significant role in the development of maritime insurance, particularly as the commercial center of Europe shifted to Antwerp in the 16th century. While many merchants may have insured to reduce their risks, insurers were largely attracted to the speculative side of the practice.³⁴ Over time, as insurance spread from its maritime context onto solid ground, it became common for policyholders to take out insurance even on cargo, lives, or events in which they had no underlying interest at all.³⁵

The line between risk-reduction and speculation in insurance has thus always been a murky one. On one hand, insurance seems to be of a piece with the risk-reducing partnerships out of which it grew, a tool for minimizing the possibility of harm by foreseeing, calculating, and transferring it. The concept of risk itself, according to historian Silvain Piron, “implicates a particular manner of relating to a future contingent event, according to an active mode of anticipation...” What most of all distinguishes this mode from others is “the fact that the

³² Barbour, “Marine Risks,” 571.

³³ Franklin notes that in 16th-century Antwerp, the center of commercial developments at the time, there was little distinction drawn in general between entrepreneurship and speculation, and that in particular insurance contracts were “used as much for speculation as for the reduction of risks.” Franklin, *Science of Conjecture*, 279. For more on the blurry line between prudence and speculation in the practice of insurance, see Clark, *Betting on Lives*, 22-23 and Barbour, “Marine Risks,” 591.

³⁴ Herman van der Wee, *The Growth of the Antwerp Market and the European Economy* (The Hague: Martinus Nijhoff, 1963), 365.

³⁵ Clark, *Betting on Lives*, 3-4, 13-21, 40-53.

anticipation leads one to think in advance about eventual consequences,” specifically the possible damages and benefits involved, such that one can then take responsibility for those consequences by offloading some of the risk.³⁶ On the other hand, however, insurance is essentially a bet on the outcome of an event between two parties whose interests are not necessarily aligned.³⁷ Even if one side insures to minimize a possible loss, the other could simply be taking his chances to secure a benefit.

Given this double-sided character and the tendencies of early traders to use insurance as a means for speculation, it would have been easy for the practice to become the sole province of gamblers. That it did not can be attributed in large part to the emergence and development of mathematical probability theory, which articulated a vision of aleatory contracts as ideal equitable partnerships. The standard of equity applied to such contracts—first by scholastic jurists and then by mathematical probabilists, or theorists of probability—combined an appreciation of their uncertain character with an attempt to render their initial conditions fair. The early probabilists thereby suggested, in effect, that wagering could be not only prudent but also just, and that justice of this kind could be attained despite the apparently arbitrary effects of chance.

³⁶ Piron, “L’apparition du resicum,” 4-5.

³⁷ See Barbour, “Marine Risks,” 591 and van der Wee, *Growth of the Antwerp Market*, 365. van der Wee notes that speculation rather than covering risks was “the most important factor for the success of maritime insurances in the course of the 16th century,” and that insurers in particular were attracted above all by the “speculative possibilities.”

II. Probability theory and the doctrine of aleatory contracts

A. The legal background of early probability theory

Mathematical probability owes both its inspiration and much of its conceptual apparatus to late-16th and early-17th century discussions regarding the equity of aleatory contracts.³⁸ The challenge for such contracts from a legal and theological perspective was how to translate uncertain future profits into fair present prices. Although the rules for doing so were primarily qualitative, they were guided by the Rule of Fellowship, which held that each party's gain should be proportional to his investment.³⁹ In the words of one early 17th-century text, a party who seeks advantage without also facing the possibility of loss does so against the "right of the association," which demands that "the chance of harm and of profit be common between the companions."⁴⁰

Where the outcome of an agreement was particularly uncertain—as in annuities, insurance, and other risky ventures—some estimation of the likely or probable gain had to be made. Out of this necessity a qualitative account of what we now know as expectation emerged, as the product of likelihood and outcome value. In principle, by equalizing this entity the parties could come to a fair agreement, ensuring that neither side exposed itself to a greater extent than was justified by its potential reward.⁴¹

³⁸ This point has been brought out most thoroughly by Daston, *Classical Probability*, ch. 1, drawing on Coumet's "La théorie du hasard."

³⁹ For a cursory survey of treatises on commercial arithmetic in 15th- and 16th-century France, including their treatment of the Rule of Fellowship, see David Murray, *Chapters on the History of Bookkeeping, Accountancy, and Commercial Arithmetic* (New York: Arno Press, 1978), 423-45. For further discussion and citations see Daston, *Classical Probability*, 20. For an overview of canon law and later scholastic treatments of this issue, see Franklin, *Science of Conjecture*, 263-69, 285-288.

⁴⁰ J. Benedicti, *La somme des péchez et le remède d'icieux... premièrement recueillie et puis nouvellement revue* (Paris: 1601), 332, quoted in Coumet, "La théorie du hasard," 591.

⁴¹ Some authors claimed that this method could render even gambling fair: "The most fundamental principle of all in gambling is simply equal conditions," wrote Jerome Cardan. "To the extent to which you depart from that equality, if it is in your opponent's favor, you are a fool, and if in your own, you are unjust." Jerome Cardan, *The Book on Games of Chance* "*Liber de Ludo Aleae*," trans. Sydney Henry Could (New York: Holt, Rinehard and Winston, 1961), 5, cited by Edith Dudley Sylla, "Introduction," in Jacob Bernoulli, *The Art of Conjecturing*, trans. Edith Dudley Sylla (Baltimore: Johns

The French jurist Jean Domat, writing at the close of the 17th century, summarized the dominant philosophical approach to aleatory contracts at the time. “In contracts in which one trades a right,” he explained,

or something else that depends on an uncertain event, and from which one can arrive either at a profit or at a loss, depending on the difference of events, one is free...to renounce all profit and discharge oneself of all loss; or to take a sum for everything one can expect to gain; or to take responsibility for a loss adjusted for all of those one has to fear. In this way a party who wishes to retire from the agreement can determine with the others what is his present and certain profit, or what he stands to lose, whatever events may arrive.... Thus they arrive between them at a kind of equality in their lots, which makes their agreement just.⁴²

Domat has here given a descriptive account of what the early probabilists tried to render mathematically exact—that is, the expectations of parties to a contract in which final profits depend on the unknown outcome. He also articulates the rationale for doing so, namely to arrive at a kind of equality between the two sides and thereby render the contract just. The fair price of each party’s lot is the present equivalent of his uncertain future winnings. Thus if one party wishes to “retire from the agreement,” preferring the security of a known sum to the possibility of a loss, he can calculate his “present and certain profit” as a function of the amount at stake and his odds of obtaining it.⁴³

Similarly, the authors of the Port Royale *Logique*, writing in 1662, explain that “to judge what one must do to obtain a benefit, or avoid a harm, one must not consider only the benefit or the harm in itself, but also the probability that it will arrive or not arrive, and regard geometrically the proportion of all of these things together.” They go on to explain, in terms even more explicit than

Hopkins, 2006), 70. Some theologians had also defended the legality of gambling in cases where there was “equality of uncertainty, peril or chance.” Daston, *Classical Probability*, 23.

⁴² Jean Domat, *Les loix civiles dans leur ordre naturel* (Paris: 1777), 30.

⁴³ Coumet points to the similarity between this statement and Pascal’s arithmetic triangle, which provided the mathematical foundation for combinatorial probability. Coumet, “La théorie du hazard,” 591.

Domat's, what it is that makes a wager or aleatory contract fair. Games in which all of the parties face exactly the same expectation are equitable, "as much as games can be," while "those which deviate from this proportion are manifestly unjust."⁴⁴ Lotteries are an obvious example of an unjust game, since those who purchase tickets pay more than their expected benefit, while the lottery master usually takes a tenth of each payment for his own advantage. Requiring that parties pay the just price or expectation value for their respective lots is thus meant to ensure that, proportionally speaking, no one stands to gain or lose more than anyone else. The earliest theorists of mathematical probability also operated explicitly within this legal and moral framework, which as a result left an indelible impression on the apparatus of insurance as well.

B. Probabilistic justice

The 1654 correspondence between Blaise Pascal and Pierre de Fermat, published in 1679 and widely regarded as the birth of mathematical probability, is a quantified version of what Domat and the *Logique* describe. In it, the two authors attempt to calculate the amount due to a player in an interrupted game of chance—someone who, as in Domat's account, prefers to take the fair price of his lot rather than test his luck by following through with the wager. Although neither Pascal nor Fermat used the term expectation, both calculated the value sought by multiplying the total amount at stake by the likelihood of the player's success in what remained of the game. As in the contemporaneous legal analyses, the working assumption is that by paying the expectation value for his lot, the departing player finds himself in a position equivalent to that of the player who continues, and the game is consequently fair.

⁴⁴ Antoine Arnauld and Pierre Nicole, *La logique, ou l'Art de penser* (Paris: 1662), 388-89.

Subsequent probabilists readily took on the same challenge—to, as Christiaan Huygens put it, “determine how much greater a share of the stakes I should get than my opponent if we agree to quit with the game unfinished”—and were even more explicit in defining the player’s expectation as the wager’s fair price. Credit for introducing the term expectation itself typically goes to Huygens, who published his *De ratiociniis in ludo aleae* in 1657 after he had heard about the Pascal-Fermat correspondence. Both Huygens and his commentator, Jacob Bernoulli, use the Latin *expectatio* and *sors* as synonyms, although *sors*, or “lot,” can mean both capital invested and expected payoff.⁴⁵ The reason for the interchangeability of the two terms is of course that in a fair game, as we have seen, one should pay or invest exactly as much as one anticipates receiving in prizes.⁴⁶

Huygens opens his treatise with the classic mathematical formula for expectation: “If I may expect either a or b and either could easily fall to my lot, then my expectation should be said to be worth $(a + b)/2$.” He then sets out to “not only demonstrate this rule, but first derive it,” and in doing so offers the first explicit proof that probabilistic expectation is the fair price of a player’s lot.⁴⁷ Huygens’ proof hinges on the device of an equivalent fair wager, for which the player would willingly pay the same amount as his expectation in the original game; because he is indifferent between the two bets, the expectation, or the amount wagered in the second, must also be the fair price of the first.⁴⁸ Unlike subsequent thinkers, Huygens establishes the fairness of the equivalent

⁴⁵ Sylla, “Introduction,” 71.

⁴⁶ Several modern commentators have pointed out that the concept of expectation used in these early probabilistic writings differs substantially from that employed later in frequentist probability theory. Unlike later probabilists, who explicitly distinguished probability as a relative frequency from outcome value, the early probabilists tended to treat expectation as the fundamental unit of analysis, a distinct entity with its own economic worth. See Daston, *Classical Probability*, 26 and Sylla, “Introduction,” 71.

⁴⁷ Bernoulli, *Art of Conjecturing*, 133.

⁴⁸ For example, in a game involving an equal chance of winning either seven or three coins, the player’s expectation is five. One can then prove that five is the fair price of this wager by imagining another fair wager in which the player would pay the same price for an equivalent lot. Huygens’s equivalent wager in this case entails two players, each of whom puts up five coins with the stipulation that whoever wins the total sum will give the loser three. With five coins,

wager not by the players' equal expectations—since this would have rendered his argument circular—but rather by the indistinguishability of their initial positions.⁴⁹ They wager the same amount, the fair price or expectation value of the first bet, and have equal chances of winning the total sum. They then all bilaterally agree to perform various symmetric trades, promising to pay each of the others a specified amount in the event that one wins. Their initial identity and reciprocity in turn ensures that in the end every player's odds of winning the total amount are the same.⁵⁰

Jacob Bernoulli, commenting on Huygens's text in his *Art of Conjecture*, published in 1713, offers a different demonstration of the fairness of mathematical expectation as a wager's price, based on what he calls “reasoning that is more popular...and more adapted to common comprehension.”⁵¹ Subtly modifying the idea of the equivalent wager, Bernoulli defines a person's expectation as “just as much as he will acquire without fail”—not after the real wager has been run, but in a kind of hypothetical alternative wager, in which the parties together claim the whole pool and then agree to divide it up proportionally among them. For example, to find the expectation of a wager in which one has an equal chance of winning a or b , Bernoulli asks us to imagine that each amount is hidden in one of two hands, and two players each choose a hand and keep its contents. In reality, of course, after the game is played, one player will receive a and the other b . Yet in Bernoulli's alternative scenario, we are asked to imagine that the total pool, $a + b$, belongs to both players in common. In that case, the two players “will acquire without fail and ought therefore to expect” the total amount hidden, or $a + b$, and this is their expectation. “But it must also be conceded that each of us has an

then, “I could again arrive at a situation in which I had an equal expectation of getting three or seven coins, contending on equal terms” with another person. Ibid. See also Hacking, *Emergence*, 96.

⁴⁹ This point is indebted to Daston, *Classical Probability*, 25-26.

⁵⁰ Thanks to the rules of the game, each has an equal likelihood of winning, and even if each player ends up with different individual chances, their sum will always equal the total pool.

⁵¹ Bernoulli, *Art of Conjecturing*, 134.

equal right to what we expect,” since the two parties’ odds of winning either a or b are identical. “Therefore it follows that the total expectation should be divided into two equal parts, and to each of us should be attributed half of the total expectation,” or $(a + b)/2$.⁵²

Thus while Huygens defines expectation as the price one would pay for an equivalent fair wager, Bernoulli defines it in terms of one’s fair portion of the total share of the bet, or the amount to which one is legitimately entitled given the initial conditions of the game. Where the players wager the same amount and stand the same chance of winning, each has a legitimate claim to an equal share—even if the amounts they ultimately win will differ. Where they wager different amounts or have different chances of success, their just claims on the total will correspondingly differ as well.⁵³ On Bernoulli’s account, then, expectation becomes the numerical expression of a kind of right, deriving from the players’ collective entitlement to the total amount and from their initial relationship to one another. The proportion between the parties’ odds and their legitimate claims on the total amount at stake is what ensures the sought-after equality between them, and with it the justice of their agreement.

C. Equipossibility and the distributive turn

Given what were undoubtedly well-known assumptions about contractual equity at the time of their writing, it is unlikely that Huygens and Bernoulli would have taken such pains simply to defend the claim that a party’s earnings should be proportional to his investment. Rather, the innovation of their new probabilistic approach *vis-a-vis* the doctrine of aleatory contracts was its claim that one can rigorously quantify each party’s lot by dividing the total pool by the number of

⁵² Ibid.

⁵³ “...[T]o the extent that one person has a greater expectation of winning, it is just that he put up that much more, if they want to play with equal lot.” Ibid., 141.

equally possible outcomes. Prior to this discovery, commentators who approached the so-called division problem arrived at such varied results that some concluded it was simply unsolvable. “[T]he resolution of such a question must be judicial, rather than mathematical,” wrote one 16th-century mathematician, “so that in whatever way the division is made there will be cause for litigation.”⁵⁴ Both Huygens’s and Bernoulli’s demonstrations are intended to show that their method does reliably and precisely generate a wager’s fair price. In both cases, the equivalent or imagined fair wager involves players who are each equally likely to win, and whose expectation is therefore an equal fraction of the total sum.

A number of modern commentators have credited Leibniz as the first to define probability in terms of equally possible cases.⁵⁵ In his *On the Estimation of Uncertainty*, dated September 1678, Leibniz sets out to justify the use of expectation as the fair price of a game. He defines a fair game as one in which there is the “same proportion of hope to fear on either side,” which is the case when “players do similar things in such a way that no distinction can be drawn between them, with the sole exception of the outcome.”⁵⁶ This is what has been called an aleatory condition—as opposed to an epistemic one—in the form of rules that result in each player’s identity, or equal chance of winning.⁵⁷ Leibniz then generalizes from this scenario to any set of equally possible

⁵⁴ N. F. Tartaglia, *General Trattato di Numeri et Misura* (Venice: 1556), I, fol. 266 1, quoted in Hacking, *Emergence*, 51.

⁵⁵ On Leibniz’s invention see Hacking, *Emergence*, 125 and Wolfgang David Cirilo de Melo and James Cussens, “Leibniz on Estimating the Uncertain,” *Leibniz Review* 14 (2004): 31.

⁵⁶ Gottfried Wilhelm von Leibniz, “On Estimating the Uncertain,” trans. Wolfgang David Cirilo de Melo and James Cussens, *Leibniz Review* 14 (2004): 43.

⁵⁷ In other words, Leibniz is not saying here that we assign equal probabilities to the players because we cannot know who will win, but rather that the game is set up in such a way that their odds of winning are objectively equal. See Cirilo de Melo and Cussens, “Leibniz,” 33.

outcomes, defining the probability of any subset as its “*aliquota portio*” or proportional part of the total.⁵⁸

Leibniz may have been the first to explicitly define probability in terms of equipossible outcomes, but one can see the foundations of Leibniz’s definition in Huygens’s fair equivalent wager. Huygens, too, understands his equivalent wager as fair because all of the players are identically situated and indistinguishable until the final outcome is revealed. Since each wagers the same sum and is equally likely to win, each one’s expectation is his proportional share of the total, and equal to everyone else’s. In thus establishing the players’ interchangeability as the main condition of his paradigmatic fair wager, Huygens anticipates Leibniz’s argument—on which his equipossibility definition of probability is based—that where all outcomes are equally likely, “my hope will be that portion of the sum of things that we get by setting it in relation to the number of outcomes.”⁵⁹

Leibniz does deserve credit, however, for introducing the language of distributive justice to the argument for a probable expectation’s fairness, which Bernoulli in turn adopted in his own demonstration and which is a key point distinguishing Bernoulli’s treatment from Huygens’s. “Let us suppose,” Leibniz writes, that in a game of chance “the whole pool pertains to all and that everyone’s hope is equal; if the players broke the game off and wanted to distribute the pool according to the hope or the claim to it, with the intention of profit, a man’s share would be owed to each one,” or in other words the total pool divided by the number of members.⁶⁰ Leibniz later clarifies that “[i]f several people share a thing, or if a thing is common to several people through the

⁵⁸ Leibniz, “On Estimating,” 46.

⁵⁹ *Ibid.*, 47.

⁶⁰ *Ibid.*, 44.

same claim, each man's claim is his share of the claim to the whole thing."⁶¹ Thus while Huygens's account of the justice between two players in a game of chance is commutative, meaning that it aims simply to equalize the financial standings of two sides of a transaction, Leibniz suggests a version of distributive justice, according to which each player's claim or desert is determined by his proportional share of a common pool.⁶²

We have already seen that Bernoulli uses this same line of reasoning to replace Huygens's fair equivalent wager proof. Pierre Montmort, in his 1708 *Essay Analyse* (published after Bernoulli had written *Ars Conjectandi* but before the latter's publication), takes a similar approach. Referring to each player's "right" over the total amount of money in play, Montmort explains that a party's claim derives from his respective probability of winning the whole.⁶³ Similarly, in determining how much two players ought to wager for a given stake, we calculate each one's proportional share of the total because "Pierre has no less of a right over [his fair portion of Paul's wager] than he has over the [amount] that he put into play."⁶⁴

This use of the language of distributive entitlement or justice in the early probabilist accounts can again be traced to the legal doctrine of aleatory contracts.⁶⁵ The earliest canon lawyers had held that the possibility of loss by both sides distinguished acceptable partnerships from

⁶¹ Ibid., 45.

⁶² The original definitions of both commutative and distributive justice are of course from Aristotle. While corrective or commutative justice requires that the parties be equalized with respect to the terms of their transaction alone, without regard for any other distinguishing characteristic, distributive justice "distributes common possessions" in geometrical proportion, in accordance with some idea of merit. While the major challenge for corrective justice is to equate quantities or goods that are unequal, like the skills of a doctor and those of a farmer, the major challenge for distributive justice is the persistence of disagreement about what kind of merit should be determinative. See Aristotle, *Nicomachean Ethics*, 1133a17, 1131a24, 1131b27, 1131b14.

⁶³ Pierre Rémond de Montmort, *Essay Analyse sur les Jeux de Hazard* (Paris: Jacque Quillau, 1708), 2.

⁶⁴ Ibid., 3.

⁶⁵ For another example of the importance of both commutative and distributive justice in scholastic analyses of contracts, see the discussion of the *mons pietatis* in Noonan, *Scholastic Analysis*, 296-299.

usurious ones, because justice required the mutual sharing of profits and losses. Such an account of justice focuses less on ensuring strict equality between two sides—the aim of commutative justice—than on equitably dividing up a shared resource. Once risk emerged as an alienable commodity, the legal challenge became how to fairly price the traded good, and as a result many juridical analyses became more explicitly commutative.⁶⁶ Yet the original understanding of a just aleatory contract as an equitable partnership persisted.⁶⁷ In fact, one might say that it is the nature of aleatory agreements to blur the line between these two understandings. On one hand, one can imagine that there is a good being traded or a service rendered whose fair price can be calculated and the parties thereby equalized. On the other hand, one can interpret the parties as intentionally banding together in the face of uncertainty, pooling their resources to cushion whatever blows may come. The distributive argument is an expression of the second view.

Leibniz and Bernoulli, in resting their own definitions of probabilistic expectation on an argument from distributive rather than commutative justice, put forward what is in fact a radical interpretation of the basis for distributive claims.⁶⁸ If the major problem for distributive justice is defining what kind of desert should be determinative, these authors implicitly solve it by identifying

⁶⁶ As Jesuit cardinal Juan de Lugo put it, “The first condition for the justice of an insurance is, that the price be equal to the peril undertaken; certainly that the price paid for the obligation should be as much as that obligation is worth in the judgment of experts.” J. de Lugo, *De Iustitia et Iure* disp. 31 sec. 7 (Lyons: 1652), 2: 447, quoted in Franklin, *Science of Conjecture*, 288.

⁶⁷ “One who contributes to a partnership 1000 ecuts, of which he wants to be reimbursed whatever fortune may arrive, and nevertheless if there is profit he wants to obtain his part, he commits a malicious usury: for he places always what is his in assurance and at the same time wants to gain against the right of the partnership, which wants that the chance of harm and of profit be common between the companions.” Benedicti, *La somme des péchez*, 332, quoted in Coumet, “La théorie du hasard,” 591.

⁶⁸ While Leibniz at times speaks the language of commutative justice—for example, in explaining at the beginning of his article that each player “buys his hope with...a fair price”—his normative argument concerns equity or proportion in divvying up a common resource. Leibniz, “On Estimating,” 43. The priority he thereby places on distributive justice may be a result of the fact that he regarded it as more authoritative than “strict” or commutative justice. See his “Meditation on the Common Concept of Justice,” in *Political Writings* (2nd ed.), ed. Patrick Riley (Cambridge and New York: Cambridge University Press, 1988), 60.

the relevant measure as each party's potential for gain or loss.⁶⁹ This potential dictates each one's contribution and gives rise to a proportional claim on the total amount at stake. In other words, it is the individual's likelihood of success or failure—independent of any other empirical fact about him and without regard to merit in any traditional sense—that determines his rightful share of the total pool.

This novel interpretation of distributive desert was enabled by the equipossibility interpretation of probability. Each equal likelihood in effect provides a baseline of entitlement against which all subsequent distinctions must be drawn. For example Bernoulli, describing the case of a player who is three times as likely as another to win a game, says that “it is clearly completely just that someone who wants to take the place of the three players...should also put up three times as much,” because distributively speaking he is the equivalent of three out of four identically situated players.⁷⁰ This is of course a fiction, designed for normative purposes, as Bernoulli implicitly acknowledges later on in explaining that “...[I]f a prince allows two criminals to contend with equal lot to live, then each of them will be judged...to have 1/2 of life and 1/2 of death. In this way a man could, even in a proper sense, be called half dead and half alive.”⁷¹ The sum wagered in a game of chance no more belongs to all of the players in common than a single life belongs to both criminals. Yet the fiction of collective ownership and of a corresponding right to one's proportional share serves the normative purpose of translating arbitrary outcomes into equitable apportionments. We cannot know in advance who will win and who will lose, and the end results will surely be divergent, but we can equalize ourselves beforehand in such a way that the outcome, no matter how differential, still seems fair.

⁶⁹ Aristotle, *Nicomachean Ethics*, 1131a24.

⁷⁰ Bernoulli, *Art of Conjecturing*, 140.

⁷¹ *Ibid.*

The doctrine of aleatory contracts, as appropriated and interpreted by mathematical probabilists, had a formative influence on thinking about insurance. While in practice many who took out insurance policies through the mid-18th century and even beyond did so with purely speculative intent, there emerged thanks to probability theory an alternative tradition of thinking about insurance as a tool for distributive justice. This tradition relied directly on the original insight that, *ex ante*, equal likelihoods give rise to equal claims. In the case of insurance, and in particular life insurance, the equiprobable outcome in question was the death of any single member of a mortality class, calculated from a growing body of government-collected statistics. The idea of deaths as equiprobable outcomes, and of similar individuals as probabilistic equals, meant that the logic outlined by Leibniz and Bernoulli could be applied well beyond games of chance, to moral and political ends.

As we will see, however, insurance, like aleatory contracts more generally, remained subject to two distinct interpretations. Insurance can be understood as a bilateral contract for the exchange of a risk, or it can be seen as a burden-sharing partnership among any number of similarly situated individuals. While it is true that commercial insurance lends itself to the first interpretation, and voluntary mutual insurance to the second, it is my contention that all types of insurance are heirs to this conceptual ambiguity, inherited from the earliest treatments of aleatory contracts and, even more deeply, from the ambiguous character of probability itself. Probability is at once an attempt to quantify good sense, or how a reasonable person would act in the face of certain odds, and a measure of empirical likelihoods, on the basis of which parties can be equalized and their distributive shares defined. The history of thinking about insurance reveals an evolving attempt to harmonize these two understandings.

III. Life insurance and probabilistic justice

A. Equity in empirical probability theory

The very first calculations of mathematical expectation were based on idealized scenarios involving coins, dice, cards, and similar games, for which a fixed number of equally possible outcomes could be identified and their relative frequencies tabulated *a priori*. Although the early probability theorists perceived the potential utility of their calculations beyond such simplified stock examples, they didn't yet have the tools for applying the calculus to fields in which only empirical or *a posteriori* likelihoods were available. Jacob Bernoulli was among the first to explicitly recognize this difficulty. How can anyone, he asked, count "the number of diseases, as if they were just as many cases, which may invade at any age the innumerable parts of the human body and which imply our death? And who can determine how much more easily one disease may kill than another...?" For this, one has no choice but to turn to observation: "What cannot be ascertained *a priori* may at least be found out *a posteriori* from the results many times observed in similar situations..."⁷²

Bernoulli's suggestion that likelihoods could be gleaned from repeated observation would inaugurate and provide a foundation for the entire theory of *a posteriori* probability. His own contribution to this theory, now known as the weak law of large numbers, showed that as the number of trials of a given phenomenon increases to infinity, it becomes increasingly likely that the difference between the true and the observed ratio of outcomes will fall within a specified interval of one's choosing. Assuming one already knows the true probability of an event, then, one can use Bernoulli's law to determine how many observations one needs in order to approach what he called moral certainty, or the degree of probability sufficient for use in practical affairs, that the observed

⁷² Ibid., 326-27.

probability will fall within a given interval surrounding the true one.⁷³ Bernoulli did not provide a mathematically rigorous way of solving the inverse problem—that is, how to infer from an observed ratio of outcomes an unknown true probability of events. Yet he did suggest a commonsense solution: “If, for example, there once existed three hundred people of the same age and body type as Titius now has, and you observed that two hundred of them died before the end of a decade...you could safely enough conclude” that Titius is twice as likely to die as to live in the next ten years.⁷⁴ Bernoulli described this way of inferring future likelihoods from past observations as what “everyone consistently does...in daily practice.”⁷⁵ The introduction of large numbers of prior observations would simply make such commonsense inductive inference more reliable.

The first treatments of *a posteriori* probability were in the domain of life insurance mathematics, where a growing body of government-collected statistics made Bernoulli-style inference from repeated observation possible. Like the early treatments of games of chance, these analyses were explicitly driven by concerns about contractual equity, and set out to calculate expectation as the fair price of an annuity or insurance contract. They also, like treatments of *a priori* probability, based all of their calculations on the equiprobable case. For example, Jan De Witt’s 1671 *Treatise on Life Annuities*, one of the first attempts to apply mathematical probability to annuities, begins with a demonstration of the formula for expectation that is strikingly similar to Huygens’s, all the way down to the fair equivalent wager.⁷⁶ In justifying the traditional formula, De Witt relies even more explicitly than Huygens on the equal likelihood of possible outcomes, and

⁷³ Despite subsequent confusion, it was actually de Moivre who figured out how to determine the number of needful observations with any degree of accuracy. See Karl Pearson, “James Bernoulli’s Theorem,” *Biometrika* 17, no. 3/4 (Dec. 1925): 205-6.

⁷⁴ Bernoulli, *Art of Conjecturing*, 327.

⁷⁵ *Ibid.*, 328.

⁷⁶ John De Witt, *Treatise on Annuities*, trans. Robbert Gibbes Barnwell, in Barnwell, *A Sketch of the Life and Times of John De Witt* (New York: Pudney and Russell, 1856), 86-88.

then shows how the same equiprobability model can be applied to mortality prospects. An annuitant, “having purchased and sunk a life annuity on a young nominee, has in possession, or in his favor, as many different expectations or chances as there are half years in which the death of the nominee may occur.”⁷⁷ If we assume that mortality remains constant in the first fifty years of life, then “the first hundred different expectations or chances...may result with the same facility, and relative to their probability are equal.”⁷⁸ After setting out a few more rough assumptions about the distribution of annuitant deaths by age, De Witt arrives at a tally of each individual’s chances of semiannual death over a lifetime, divides the total potential annuity payment by this number, and arrives at the contract’s expected value.

It is almost certainly not an accident that the first life insurance mathematicians chose to express mortality prospects in terms of equally likely cases.⁷⁹ Like De Witt, they did this by basing their calculations on mortality curves, which allowed them to regard the chance of annual death as equally likely through the prime of life, and as some multiple of that baseline in other periods. This assumption allowed them to calculate expectation using the established—and, at the time, only available—method of simply dividing the total amount at stake by the number of equal chances of “winning” it. Since the goal of contractual equity was as much a motivating concern for the life insurance mathematicians as for their *a priorist* contemporaries, it is easy to understand why they would have looked for a way to translate mortalities into equal probabilities.

⁷⁷ Ibid., 99.

⁷⁸ Ibid.

⁷⁹ Ian Hacking has suggested that the use of the equiprobability model in *a posteriori* probability may have been little more than an accident, due to the fact that the early life insurance mathematicians chose to base their calculations on roughly uniform mortality distributions. Hacking, *Emergence*, 121. This may be true, but only in the sense that the first examples of life insurance mathematics happened to emerge directly out of probabilistic treatments of games of chance, and were therefore heavily influenced by the latter’s normative concerns and analytic framework.

De Witt's work, which was in fact a report to the States General of Holland advising the government on how to price its annuities, ultimately proposed a single annuity price equal to the total sum that the government could anticipate paying divided by the number of annuitants.⁸⁰ The price therefore reflected the government's expectation, but not the particular expectations of annuitants of different ages, who entered the contract with different life expectancies depending on their age of purchase. In this De Witt was not unique—most annuity schemes throughout the 17th century priced their products without regard to age.⁸¹ But as a consequence, his analysis had obvious flaws from the point of view of contractual equity. Most annuitants facing such an average price would reasonably select healthy children as their nominees, giving them a considerable advantage in the contract (as De Witt himself was quick to point out).⁸²

It was not until Edmund Halley published his seminal table of deaths by age and used the results to calculate annuity values that expectation as a function of individual life expectancy began to make its way into life insurance mathematics. Halley's report, presented to the Royal Society of London in 1693, notes that the purchaser of an annuity "ought to pay only such a part of the value of the *Annuity*, as he has Chances that he is living; and this ought to be computed yearly, and the Sum of all those yearly values being added together" equals the total value of the annuity.⁸³ In other words, it is only fair to charge an annuitant for the annuity's expected value *to him*, which is its

⁸⁰ De Witt, *Treatise*, 96-102. For analysis, see Anders Hald, *A History of Probability and Statistics and their Applications before 1750* (New York: Wiley, 1990), 123-25.

⁸¹ Hald, *A History of Probability and Statistics*, 118. The city of Amsterdam did, in 1672, take what was then the remarkable step of offering annuities at a price dependent on the annuitant's age, although their price was extremely low compared with the results obtained by De Witt and other mathematicians at the time. Even after the 1694 publication of Halley's life table and corresponding calculations, the English government continued selling annuities at a price independent of the annuitant's age, to the purchasers' advantage. See also *Ibid.*, 131, 139.

⁸² *Ibid.*, 118 and De Witt, *Treatise*, 104.

⁸³ E. Halley, "An Estimate of the Degrees of the Mortality of Mankind, Drawn from Curious Tables of the Births and Funerals at the City of Breslaw; With an Attempt to Ascertain the Price of Annuities upon Lives," *Philosophical Transactions* 17 (1693): 602.

present value multiplied by his odds of living through each subsequent year. Equity in *a posteriori* probability was thereby rendered consistent with the earliest *a priori* calculations.

It was the use of mortality tables that rendered the *a posteriori* interpretation of probabilities comparable to that of the *a priori* accounts. Rather than referring to the individual's equal chance of dying each year or half year over the course of a certain period, as in De Witt's analysis, equiprobability from Halley onward referred to the individual's equal chance of dying *relative to everyone else his age*. As Halley put it, "as the number of Persons living after that term of years, to the number dead; so are the Odds that any one Person [that age] is Alive or Dead." This means that just as in Huygens's and Bernoulli's fair wagers, each party can legitimately claim an equal fraction of the total sum because each has an equal potential for "winning" or "losing" the wager. As an analytical tool, then, mortality tables were uniquely well suited to promoting equity in life insurance mathematics according to the dominant approach of the time. The fact that they became the method of choice for calculating expectations confirms the priority of normative considerations—specifically, the need for a baseline of equality against which distributive claims could be measured—in *a posteriori* just as in *a priori* probability theory.

B. The birth of statistical life insurance

In practice, it took over half a century before the insights of mathematical probability found their way into life insurance contracts. While some early insurance cooperatives and commercial providers did make rudimentary use of demographic data, they did so by simply limiting their coverage to groups of people who shared roughly the same traits, such those in the prime of life or

with the same occupation.⁸⁴ Like De Witt and other early probabilists, many life insurers regarded the chance of death as roughly constant between adolescence and midlife, leading them to disregard the need for a breakdown of mortality risk by age.⁸⁵ This meant that their schemes were limited in scope and met with wildly varying success.⁸⁶

It was not until the final third of the 18th century that an insurance company used statistical probabilities to set its premiums, and as a result was able to accommodate a wider range of customers than any one company or cooperative had previously been able to do.⁸⁷ The Society of Equitable Assurances on Lives and Survivorships, or Equitable, received its founding impetus from mathematician James Dodson. A student of Abraham De Moivre, one of the first mathematicians to turn mortality statistics into probabilities, Dodson had been denied admission to London's Amicable Society on the ground that at age 45 he was too old for coverage. As a result he set out to establish a life insurance association that would grade premiums on the basis of age and therefore be able to accept older customers without jeopardizing its viability.⁸⁸ The preamble to the Equitable's deed of settlement speaks of "great benefits" to be derived from life insurance based on "premiums

⁸⁴ The most sophisticated of these was probably the Scottish Widows' fund, discussed and in later editions praised by Richard Price in his *Observations on Reversionary Payments*, 6th ed. (London: T. Cadell and W. Davies, 1803), 108-115. See also Clark, *Betting on Lives*, 121-25.

⁸⁵ Clark, *Betting on Lives*, 122-23.

⁸⁶ For example, the Royal Exchange and London Assurance offered a flat rate but limited their coverage to short-term policies. They also charged a higher premium to those undertaking very long journeys and excluded customers with smallpox. Daston, *Classical Probability*, 170-71. The Amicable Society offered long-term life insurance, but its model of fixed payments and equal dividends required it to limit admission to men in good health between ages 12 and 45 and to cap its membership at 2,000. Most of the later widows' funds did scale members' payments to age, but they fixed those payments primarily by guesswork and did not consider that claims on the fund would increase as their member population aged. As a result they failed at an alarming rate. *Ibid.*, 169. Clark contrasts the "spectacularly disastrous results" of Assheton's Annuity Scheme, which lacked demographic knowledge, with the Church of Scotland's scheme, which carefully traced the habits and mortality experience of its members, and was thus able to anticipate the cost of its reversionary annuities and change its fees accordingly. Clark, *Betting on Lives*, 134-46.

⁸⁷ Daston, *Classical Probability*, 174-75; see also A. Fingland Jack, *An Introduction to the History of Life Assurance* (London: P.S. King and Son, 1912), 233.

⁸⁸ See Augustus De Morgan, "Some Account of James Dodson, F.R.S.," *Journal of the Institute of Actuaries and Assurance Magazine* 14, no. 5 (Oct. 1868): 346.

proportionate to the chance of death attending the age of life to be assured, and to the time such assurance is to continue.”⁸⁹

Given the importance of contractual equity to probability theory, including early life insurance mathematics, one needn’t look far to understand why Dodson chose to call his company the Society of *Equitable Assurances*.⁹⁰ By placing individuals into groups within which their expectations could be considered equal, age-graded life insurance sought to avoid the perceived unfairness of prior plans. Calculating prices on the basis of average mortality by age would ensure both that individual subscribers were probabilistically equal with one another at the commencement of the contract, and that the society itself would have enough funds to pay those who lived the longest.⁹¹ William Morgan, one of the Equitable’s first actuaries, explained the approach as follows: Assume, based on mortality data of the time, that one out of every fifty 39-year-olds in London will die before reaching the age of 40. If fifty Londoners each purchased a life-insurance policy promising \$100 in the event of death within the year, that policy would cost them each 1/50 of the

⁸⁹ The Society for Equitable Assurances, *A Short Account of the Society for Equitable Assurances on Lives and Survivorships* (London: 1762), 6, 19.

⁹⁰ Richard Price, who served as actuarial advisor to the Equitable, critiqued the plan of London’s Amicable Society as inequitable for a variety of reasons, among them its “requiring the same payments from all persons under 45, without regarding the differences of their ages; whereas, the annual payments of a person admitted at 45, ought to be double the annual payment of a person admitted at 12.” Price, *Observations*, 164. By contrast, Price notes that the Equitable “assures any sums or reversionary annuities on any lives, for any number of years as well as for the whole continuance of the lives, at rates settled by particular calculation; and in any manner that may be best adapted to the views of the person assured. That is, either by making the assured sums payable *certainly* at the failure of any given lives, or on condition of survivorship: and also either by taking the price of the assurance in *one present payment*, or in *annual payments* during any single or joint lives, or any terms less than the whole continuance of the lives.” *Ibid.*, 176. William Morgan, another of the Equitable’s early advisors, notes of the Amicable’s approach that “[n]othing could exceed the injustice and improvidence of such a plan, which made no distinction between the old and the young in its premiums...” William Morgan, *A View of the Rise and Progress of the Equitable Society* (London: Longman, Rees, Orme, Brown, and Green, 1828), 14.

⁹¹ If Morgan’s account is to be trusted, however, it took several years before the Equitable attempted to manage its finances with any mathematical rigor. See Morgan, *Equitable Society*, 16-19.

total value, or \$2. The insurance company would thus collect \$100 in total and redistribute it to the one who happened to pass away.⁹²

As in the wagers of *a priori* probability, then, the principle supporting statistical life insurance schemes is that those who end up “losing” the bet about their own longevity compensate those who “win.”⁹³ Here the psychological impetus for participation is even clearer, however, since it directly reflects the burden-sharing logic of aleatory contracts. The contract begins with the assumption that the insured parties are interchangeable, since their individual mortality is the average for the class and thus identical across the group.⁹⁴ Because they all face the same odds of death, they agree in advance to use their collective resources to compensate one another for losses any one of them incurs. In presenting itself as a society rather than a company, the Equitable explicitly set out to reflect this image of mutuality.⁹⁵ Once each insured party entered into the group, he in a sense became a common owner of the total pool. Indeed, after running a surplus for years due to its conservative mortality estimates and other factors, the Equitable began to regularly redistribute some of its excess funds to members in rough proportion to the values of their policies.⁹⁶

⁹² William Morgan, *The Principles and Doctrine of Assurances, Annuities on Lives, and Contingent Reversions, Stated and Explained* (London: Longman, Hurst, Rees, Orme, and Brown, 1821), 2.

⁹³ Dodson himself explicitly analogized life insurance to games of chance in his 1756 “First Lecture on Insurances” and explained that the method of calculating probabilities is the same in both. James Dodson, “First Lecture on Insurances,” 11-12, reproduced from handwritten copies transcribed from the original and available with permission of Thomas G. Kabele, <https://www.actuaries.org.uk/learn-and-develop/research-and-knowledge/library-services/historical-collections/archive-equitable-1>.

⁹⁴ On this point, see also Richard Price, “Observations on the Expectations of Lives, the Increase of Mankind, the Influence of Great Towns on Population, and Particularly the State of London with Respect to Healthfulness and Number of Inhabitants. In a Letter from Mr. Richard Price, F. R. S. to Benjamin Franklin, Esq; LL.D. and F. R. S.,” *Philosophical Transactions* 59 (1769): 90-93.

⁹⁵ Whether the Equitable was always faithful to this image is a different question, and beyond the scope of my discussion here. The difference between insurance as a commercial undertaking, in which the insurer profits from the arrangement, and non-profit mutual assurance, is an important one that generated considerable analysis at the time. See e.g., Pierre-Simon Laplace, *Philosophical Essay on Probabilities*, trans. Andrew I. Dale (New York: Springer-Verlag, 1994), 89.

⁹⁶ See Morgan, *Equitable Society*, 53-60. Morgan notes, however, that doing so with “perfect justice” was impossible, as it would have required significant feats of information-gathering and calculation each time a division took place.

The idea of a risk-spreading partnership always assumed that there is safety in numbers, even if the partners were only two instead of one. What was novel in statistical life insurance was its reliance on the numerical average, and its resulting awareness that security requires not just a few partners but a critical mass of participants who can, thanks to the notion of equiprobability, be considered interchangeable with one another. Without such a mass—say, the fifty 39-year-olds in the example above—there is no reason for anyone to enter into the contract, since each would have to pay far more than the expected value of his coverage. Even if the minimum mass is attained, however, the actual mortality experience of the group may still vary considerably from the average derived from mortality tables. As Bernoulli’s weak law of large numbers had shown, the larger the sample, the closer the observed frequency will come to the “true” one. The Equitable’s prospectus reflected this assumption in stressing that life expectancy, while uncertain in the individual case, “in an aggregate of lives is reducible to a certainty.”⁹⁷ In order for the finances of insurance to work, at least in theory, the group of insured has to manifest an average mortality that approaches the one on which their contributions are based, and this only becomes likely as their number grows.

Bernoulli’s original insight and its subsequent iterations became the mantra of a new approach to protecting the individual from any number of dangers. The promise of predictability entailed in the statistical average, in which short-term deviations give way to long-term regularity, seemed to hold the key to security and peace of mind even in the face of life’s most frightening realities. We cannot know what will happen to any individual, goes the logic, but we can know with mathematical precision what will happen in the aggregate, and we can take responsibility for our fates in light of that knowledge. Premium life insurance was the pioneer and poster child of this argument, since it was the first and for some time the only practice to consistently make use of

⁹⁷ Equitable, *A Short Account*, 3.

statistical averages.⁹⁸ With it, a new interpretation of risk began to emerge. Where early marine insurers had taken a prudential, particularized approach to quantifying a ship's likelihood of success, life insurance underwriters now relied on aggregation and abstraction.⁹⁹ Once individuals were grouped together by age, they were assumed to be alike in every other relevant respect, such that their chances of dying at a particular time could be considered equal.¹⁰⁰ With the large-numbers approach to probabilities, in other words, risk became social, both in its definition and in its use: Only by placing the individual within a group of probabilistic equals could his own expectation be calculated, and only by joining together with enough of his peers could that expectation be used as a tool for individual security.

C. Mutualism with and without risk

Probability theory thus offered an image of distributive justice as tied to rigorously calculated and largely impersonal risk assessments. Risk on this account belonged to the individual, who could transfer it to an insurer or mutual society for a suitable premium. Yet statistical risks also abstracted considerably from the individual's particularities, focusing on a small number of his features and disregarding others. This abstraction is what allowed for the calculation of probabilities from

⁹⁸ While fire insurance had existed in England since the 1680s, companies failed to analyze their own underwriting data or share data among themselves until the beginning of the 19th century. As a result they could not make precise measurements of the differential rates of fire in different buildings, or indeed place their pricing on any kind of rigorous empirical footing. Robin Pearson, "Fire, Property Insurance, and Perceptions of Risk in Eighteenth-Century Britain," in *The Appeal of Insurance*, ed. Geoffrey Clark et al. (Toronto and Buffalo, NY: University of Toronto Press, 2010), 97-98. Marine insurers did raise or lower premiums depending on certain distinguishing factors, but their attention to the particular risks of each sail actually led them away from a statistical approach to underwriting. Clark, *Betting on Lives*, 7. Ultimately, it was the financial success of the Equitable and its successors that fully established the idea that aggregate regularities are the key to individual security. Daston, *Classical Probability*, 174-75.

⁹⁹ Daston, *Classical Probability*, 120.

¹⁰⁰ In fact early English insurers ignored nearly all other distinguishing features, including occupation, health, and even sex, in setting their premiums. Timothy L. Alborn, *Regulated Lives: Life Insurance and British Society, 1800-1914* (Toronto and Buffalo, NY: University of Toronto Press, 2009), 116.

empirical observation, since only by isolating features such as sex, age, and location was it possible to generate sufficient data and with them reliable averages.

The abstract and impersonal character of such probabilistic justice stood in contrast to another popular model of mutual insurance: that of the friendly societies, or voluntary mutual assurance associations. Friendlies had first emerged in England as local, voluntary groups that brought together neighbors or members of the same trade for socializing and the pledge of assistance in the case of need. From their inception, England's societies were voluntary, and they remained local both in scale and in character through the early 1800s.¹⁰¹ Some scholars have traced their origin to trade guilds, which in the words of one 19th-century treatise “were originally instituted by the mutual agreement of friends and neighbors, and had no further object than the relief of brethren in times of distress,” though they too involved a great deal of “feasting and conviviality” among their tight-knit members.¹⁰² Whatever the friendlies' precise origin, it is generally agreed that their modern variety dates to the late-17th or early-18th century, and that their number began to grow in the early 1700s.

Formed among men who knew each other or were in some sense part of the same community, the friendlies usually charged equal fees and paid a flat benefit to any member who encountered a situation of need. They also used a portion of the dues collected for club meetings, which members saw as central to the overall benefits of belonging. In fact, these lively gatherings were so important to the ethos of the friendlies that many retained the character of a local club, where members came to drink and socialize along with pledging their mutual aid, well into the 19th century.¹⁰³

¹⁰¹ P. H. J. H. Gosden, *The Friendly Societies in England, 1815-1875* (Manchester: Manchester University Press, 1961), 17.

¹⁰² Charles Ansell, *A Treatise on Friendly Societies* (London: Baldwin and Cradock, 1835), 9, 10.

¹⁰³ Geoffrey Finlayson, *Citizen, State and Social Welfare in Britain, 1830-1990* (Oxford: Clarendon Press, 1994), 41-42.

True to their name, then, the friendly societies were guided more by norms of neighborliness and camaraderie than by those of market transfer or contractual justice. The element of financial self-interest, while clearly present, was bound up with social ties and communal aspirations that prioritized members' equal standing over individualized pricing and benefits. Yet the flip side of this egalitarian ethos was a demand for homogeneity and behavioral compliance, lest any member end up imposing a disproportionate cost on the group.¹⁰⁴ For example, the societies typically restricted their membership to men (and not women) from the same village or trade, within a limited age range, and without any known propensities to become ill or disabled. They also monitored their ranks to make sure no one was engaging in the kind of behavior that could result in excessive claims.¹⁰⁵ The friendly societies thus consciously ensured uniformity through admissions and behavioral standards that not only promoted social cohesiveness, but also justified the equal financial burden each member was expected to shoulder.¹⁰⁶

The friendlies therefore represent an alternative model of insurance as more personalized and communal than probabilistic life insurance. Rather than understanding mutual insurance as the transfer of a quantified risk, they grouped together people who were roughly similar, who knew each other well, and who could police one another's compliance with norms of responsible conduct. While probabilistic insurance abstracted from individual particularities to form its groups of burden-sharing equals, mutual societies could rest on a much thicker form of identification or affinity, one rooted in a lived sense among members that they shared a similar fate.

Thick identity had its downsides, however. Because the friendlies achieved their egalitarian

¹⁰⁴ Clark, *Betting on Lives*, 124-25; Simon Cordery, *British Friendly Societies, 1750-1914* (Hampshire: Palgrave Macmillan, 2003), 25-6; Finlayson, 41-42.

¹⁰⁵ See Frederick Morton Eden, *Observations on Friendly Societies for the Maintenance of the Industrious Classes during Sickness, Infirmary, Old Age, and other Exigencies* (London: J. White and J. Wright, 1801), 14-15 and Daniel Gottlieb, "Asymmetric Information in Late 19th Century Cooperative Insurance Societies," *Explorations in Economic History* 44 (2007): 271.

¹⁰⁶ Clark, *Betting on Lives*, 124-25.

aims with a number of rather exclusive practices, the majority of participants came from the better-off sections of the working classes, even at the height of the movement's popularity.¹⁰⁷ From this perspective, one advantage of probabilistic insurance is that it is impersonal and impartial, dependent on calculation rather than fellow feeling or the fallible people who enforce it. The potential scope of its application is also much greater, as statistical risk classes can transcend localities and affective ties. These indeed were among the arguments for the movement to reform friendly societies along actuarial lines—that is, to align members' payments with their personal expectations or risks—which began in the late-18th century and took direct inspiration from the success of the Equitable and similar companies.

The same considerations also motivated the original idea of social insurance. Publicly orchestrated insurance schemes promised to combine the mathematical rigor of statistical insurance with the communal spirit of the friendlies. They would spread a kind of bourgeois virtue to the lower reaches of the working class—those often excluded from voluntary mutual aid—by encouraging foresight, savings, and a sense of familial duty. They would also replace the friendly societies' isolated pockets of fraternity with the more encompassing and equalizing protection of the state. As the next chapter will argue, a unique constellation of probabilistic ideas supported this vision, and its practical aims in turn found expression within the probability calculus itself.

¹⁰⁷ Bernard Harris, *The Origins of the British Welfare State: Social Welfare in England and Wales, 1800-1945* (Hampshire and New York: Palgrave MacMillan, 2004), 84. Cordery disputes the claim that only well-paid and skilled workers joined friendly societies, but also emphasizes that friendlies consciously excluded workers in dangerous or unhealthy occupations and tightened their membership qualifications as actuarial science improved. He also notes that while people in low-status and poorly paid jobs joined various societies, mid-Victorians “did mentally divide themselves into two segments, an upper echelon of affiliated orders with relatively high subscription rates and a subordinate group of neighborhood clubs and burial societies,” which provided only funeral benefits. We therefore need not subscribe to the “labor-aristocracy” thesis in order to acknowledge that the reach of friendly societies' coverage was relatively limited. Cordery, *British Friendly Societies*, 72; cf. 3, 26.

Chapter Two: Probabilistic Justice and the Beginnings of Social Insurance

As we saw in the last chapter, the statistical turn in mathematical probability enabled life insurance to take off as an industry, allowing advocates to downplay the speculative or wager-like aspect of the contract and cast it instead as a fair risk-reducing partnership. From the time of the Equitable's founding, life insurance was marketed to the responsible bourgeois as a way to provide for his family's long-term welfare by tapping into the certainty of aggregate regularities.¹⁰⁸ Such mutualistic insurance, founded on the impersonal identity of equal probabilities, could at once promote both individual welfare and distributive justice. In the words of the Equitable's prospectus, "the assured being at the same time mutually assurers one to the other, the interest of one might be the interest of both."¹⁰⁹

The ideal of a union between self-interest and distributive justice also inspired both the friendly-society reform movement and the first proposals for social insurance in the late 18th and early 19th centuries. Probabilistic insurance promised to replace vulnerability and dependence with foresight and self-sufficiency. In the two sections that follow this one, I will consider in detail the mathematical and moral arguments that supported the push for friendly-society reform and state-run insurance schemes during this period. First, however, we will see how actuarial mutual insurance—risk pooling among probabilistic equals, with contributions and benefits determined by mathematical expectation—came to be a political tool and not only a means for commercial benefit or voluntary mutual aid.

The first part of the discussion focuses on Great Britain, where both life insurance and friendly societies have a long history and where the movement to reform the latter on actuarial lines was particularly prominent and successful. The second part looks to contemporaneous proposals to

¹⁰⁸ See Daston, *Classical Probability*, 177-79.

¹⁰⁹ Equitable, *A Short Account*, 18.

implement actuarial mutual insurance on a national scale, which originated in both England and France but had a particularly strong impact in the latter. Both efforts set out to bring order and mathematical discipline to the allegedly haphazard affairs of the working classes, and both did so by understanding risk as precisely calculable for large groups of people who share salient features. The origins of social insurance can thus be found in a newly socialized understanding of risk, enabled and encouraged by the increasingly statistical cast of probability theory.

I. Precursors to social insurance

A. Friendly society reform: social insurance writ small

In the last chapter, I emphasized the distinctions between friendly societies and actuarial insurance, arguing that the former represented a model of mutual aid that called on strong feelings of identity among members who were expected to share one another's burdens. Such mutual aid also, I pointed out, required a fair amount of exclusivity and behavioral compliance, which limited the friendlies' reach among the weaker segments of the working class.

The movement to reform friendly societies along actuarial lines grew from a recognition of their salutary effects and a critique of their alleged failures. This movement, which bore its first fruits in England toward the end of the 18th century and subsequently took root in France, was inspired by the success of the Equitable and the new image of life insurance it fostered. Essential to this image was the reliability of probability values derived from mortality data.¹¹⁰ Once mathematical calculations could be relied upon to accurately predict the experience of subscribers, providers could in principle offer annuities and premium insurance without taking on greater risks than they were

¹¹⁰ Daston, *Classical Probability*, ch. 3, esp. 182-83, and Clark, *Betting on Lives*, ch. 4, esp. 144-46. For a characteristic 19th-century account of the difference between probabilistic life insurance and speculation, see the discussion in Morgan, *Doctrine of Assurances*, xi.

able to financially manage.¹¹¹ So too could friendly societies accommodate greater and more diverse membership without compromising their fairness or financial viability.

Among the pioneers of this movement was Richard Price, who had served as an early mathematical advisor to the Equitable and whose nephew and protégé, William Morgan, later became its actuary. Price, whose support for the French Revolution would later make him the target of Edmond Burke's *Reflections on the Revolution in France*, had been admitted to the Royal Society in 1765 for his research on probability. In 1771, he published an influential reference work, *Observations on Reversionary Payments*, which laid out actuarial calculations for premium life insurance and annuities with the aim of offering “direction and assistance” should existing friendly societies “determine to reform themselves, or should any institutions of the same kind be hereafter established.”¹¹²

Price also helped to devise a 1773 bill that would have allowed parishes in England and Wales to sell life annuities at probabilistically determined prices and, where necessary, charge them to the parish as a security to poor purchasers. Price had supplied the tables from which the annuity prices were to be calculated, and although both the 1773 proposal and a subsequent bill of 1789 met with defeat in the House of Lords, a successful reform act of 1793 rendered Price's calculations the standard reference for friendly society contributions until 1825.¹¹³ This bill, known as the Rose Act, recognized a right to form mutual aid associations and required such groups to submit to judicial regulation to ensure their practices were financially sound. In 1803, another bill determined which

¹¹¹ This also helps to explain the success of the Church of Scotland's widely praised annuity scheme, which was the most successful of the early life insurance societies. The Church benefited significantly from the ease of collecting and analyzing information about its members, and as a result was able to use mortality data in setting prices. Yet while remarkably effective, this model did not make significant inroads into the life insurance market in the way that the Equitable's did, perhaps because its circumstances were unique and difficult to replicate by other societies at the time. See Clark, *Betting on Lives*, 146.

¹¹² Price, *Observations*, xviii-xix.

¹¹³ Maurice Edward Ogborn, *Equitable Assurances* (London: George Allen and Unwin, 1962), 96; Ansell, *Treatise on Friendly Societies*, 12.

magistrates would be charged with examining the societies' rules, and in 1817 the friendlies were permitted to deposit their capital into savings accounts.

As the advantages afforded to friendly societies grew, so did their number: While in 1802, England had only 9,672 individuals participating in such societies, by 1815 there were 925,429.¹¹⁴ Despite their growing popularity, however, many societies were still relatively unstable, creating, in the words of French social-insurance advocate Emil de Girardin, “inextricable financial embarrassments.”¹¹⁵ Parliament, trying to remedy matters, passed a new act in 1819 giving judges several new rights, among them the power to decline authorization to societies in small districts or districts where there was already another group organized on the same basis; to decline societies whose financial components were not approved by two professional mathematicians; and to ensure that no society would be dissolved until two actuaries decided that doing so would be advantageous to all its members. Additional acts passed in 1829 and 1846 recognized a right of all citizens to create mutual societies against any eventuality that could fall under the probability calculus and established the foundations of a centralized system in the form of an official registration process.¹¹⁶ As a result of such official state encouragement, there were over two million Englishmen participating in friendly societies by 1872.¹¹⁷

Given the direction that friendly society reform took, it is difficult to deny the influence of mathematical probability and actuarial life insurance on its development. With the help of proper tools, the thinking went, friendly societies, like life insurers, could encourage essential virtues such as

¹¹⁴ “Report of the Select Committee of the House of Commons,” 1825, reproduced in Ansell, *Treatise on Friendly Societies*, 17. See also Harris, *Origins of the British Welfare State*, 81.

¹¹⁵ Émile de Girardin, *La politique universelle*, 3rd ed. (Paris: 1855), 369.

¹¹⁶ Cordery, *British Friendly Societies*, 86.

¹¹⁷ Harris, *Origins of the British Welfare State*, 81.

self-reliance, frugality, and responsibility.¹¹⁸ As for their friendliness, the rough and implicit justice of the early societies was to give way, under the influence of these reforms, to the more exacting mutualism of disaggregated groupings. One relatively early pamphlet criticized existing societies for excluding members above age 35, and argued that the “greater hazard” posed by older members could be accommodated by simply grading admissions fees or periodical payments by age, as in “the practice of insurance offices.”¹¹⁹ Another popular treatise, from the height of the reform movement, cited the “manifest...injustice” of “requiring men of different ages to pay a like rate,” and found the practice at odds with the feeling of “equity and benevolence” on which the friendlies rested.¹²⁰ Its author, Charles Ansell, who also inaugurated the actuarial approach to sickness in friendly societies, emphasized the importance of basing mortality projections on “persons placed in like circumstances, except as to age,” in order to match the observed experience with the projected one.¹²¹ The implications of this reasoning are clear: The more precisely one can classify observations into risk groups without sacrificing their number, the more accurate a society’s projections and the more equitable its benefit scheme would be.

Reformers apparently saw no conflict between risk segregation and the friendly societies’ mission of communal aid because on their view actuarial insurance is an inherently mutualistic practice. Probabilistic equals who survive longer than the average pay for those who, by no fault of their own, are afflicted sooner. It is also inherently communal, in that only by joining together with a large group of peers can one ensure that the projected average is realized. Self-interest thus aligns

¹¹⁸ See Eden, *Observations on Friendly Societies*, 10; Ansell, *Treatise on Friendly Societies*, 1-2; and the discussion in Daston, *Classical Probability*, 174-78.

¹¹⁹ Eden, *Observations on Friendly Societies*, 14-15.

¹²⁰ Ansell, *Treatise on Friendly Societies*, 106-7.

¹²¹ *Ibid.*, 48. On the movement to introduce actuarial calculations for sickness, see James C. Riley, “Disease without Death: New Sources for a History of Sickness,” *Journal of Interdisciplinary History* 17, no. 3 (1987): 554.

with the common good, as the simple act of saving a small portion of one's earnings guarantees security to both oneself and the association.

As a result, it became possible during this period to present mutual insurance as an obligation, a sign of responsibility and self-sufficiency, and even civic duty. In 1825, a select committee of the House of Commons declared mutual insurance preferable to private savings, since the individual who saves "is really the speculator," betting that he will stay healthy during his productive years and live only long enough to exhaust what he has saved. The insured, by contrast, while reducing his own monetary reserves, guarantees his and others' security:

[W]henver there is a contingency, the cheapest way of providing against it is by uniting with others, so that each may subject himself to a small deprivation, in order that no man may be subjected to a great loss. He, upon whom the contingency does not fall, does not get his money back again, nor does he get for it any tangible benefit; but he obtains security against ruin and consequent peace of mind. He, upon whom the contingency does fall, gets all that those, whom fortune has exempted from it, have lost in hard money, and is thus able to sustain an event which would otherwise overwhelm him.¹²²

According to this rationale, it is not enough to simply have foresight and look after one's own good. Responsibility means tapping into the power of the average so that one's exposure is buffered by the reliable experience of the group.

Mutual insurance is not only safer but also nobler than private savings, reformers argued, combining economic good sense with a kind of fellow feeling or generosity. "The principle of the savings bank is that every man is to save for himself," explained one pamphleteer. "[T]he principle

¹²² "Report of the Select Committee of the House of Commons," 1825, reproduced in Ansell, *Treatise on Friendly Societies*, 21.

of the Friendly Societies is, that every man is to save for himself, *if he needs it*; but if not, for those whose necessities may be greater than his own.”¹²³

All of these arguments for actuarially reformed friendly societies—their economic and psychological benefit to the individual; their advantages for the group and for the public as a whole; their expression of a kind of rational altruism—found clear echoes in early arguments for social insurance. The latter were equally inspired by the probabilistic account of justice, imagining the state as one large mutual insurance society. If the original friendlies reflected a voluntary, self-policed, and therefore limited form of risk sharing, their reformed cousins showed how the paternalistic power of the state and the impersonal dictates of mathematics could rationalize self-help for the benefit of all. In this sense, the reformed friendly societies can be seen as social insurance writ small, and it was not a far stretch to propose state-run mutual insurance based on the same principles.¹²⁴

B. The first social insurance plans: mutual insurance writ large

The first proposal for something like social insurance in this vein may have come from Daniel Defoe, who in his 1697 *An Essay Upon Projects* envisioned a wide-ranging, borderline-compulsory system of publicly administered local insurance entities, including friendly societies for sailors and widows and local pension offices for laborers.¹²⁵ Defoe was clearly aware of the relevance of probability theory to his proposals and refers twice to William Petty’s *Political Arithmetic*,

¹²³ J. W. Cunningham, *A Few Observations on Friendly Societies and their Influence on Public Morals* (London: 1817), 14 (emphasis in original).

¹²⁴ For instance, one early 19th-century parliamentarian, John Christian Curwen, advocated a National Benefit Society that would be administered by vestries with equal funding from taxpayers, workers, and employers on the same principles that guided his strong support for friendly societies. Such institutions would create “respectable members of society” who would in turn take responsibility for their own wellbeing and reduce expenditures on poor relief. J. C. Curwen, *Hints on Agricultural Subjects* (London: J. Johnson, 1809), quoted in Cordery, *British Friendly Societies*, 48-9.

¹²⁵ Daniel Defoe, *An Essay Upon Projects* (1697), Kindle edition, 55-56.

though he does not reveal a deep engagement with the mathematical literature. His schemes were to be universal in scope, yet he insisted that fairness requires disaggregating individuals into appropriate risk categories.¹²⁶ Since “a seaman or a soldier are subject to more contingent hazards than other men,” they are “not upon equal terms to form such a society; nor is an annuity on the life of such a man worth so much as it is upon other men.”¹²⁷ It is therefore “necessary to sort the world into parcels,” and as men’s “contingencies differ, every different sort may be a society upon even terms.”¹²⁸

Defoe recognized that such mutual aid differed from the premium insurance offered by commercial offices at his time. Although he emphasized that payment for the acceptance of risk was a legitimate commercial custom, he noted that in practice it often devolved into gambling. Mutual insurance, by contrast, where “all who subscribe pay their quota,” could serve the public interest without such speculative pitfalls.¹²⁹

Defoe’s work fell largely on deaf ears at the time, but found resonance with later friendly society reformers.¹³⁰ By then, the success of statistical life insurance had demonstrated that disaggregation and differential pricing could allow mutual societies to achieve greater inclusiveness without sacrificing their financial stability. What is more, because the theory underlying such insurance stressed its fairness and non-speculative character, there was no reason why its model should be applied only to commercial insurance. In fact, if anything it seemed better suited to cooperative welfare schemes, which aimed not to make a profit but to offer security through

¹²⁶ Ibid., 47.

¹²⁷ Should a society be formed on such unequal terms, “the seaman’s executors would most certainly have an advantage, and receive more than they pay.” Ibid.

¹²⁸ Ibid.

¹²⁹ Ibid., 44-46, 64.

¹³⁰ See, e.g., William Harral Johnson, *The Past, Present, and Future of Friendly Societies* (London: F. Farrar, 1867), 8.

communal self-help.

Nearly a century after *An Essay Upon Projects* was published, and at roughly the same time that Richard Price was campaigning on behalf of the Rose Act, English clergyman John Acland used Price's data to propose a plan for national insurance similar in spirit to Defoe's. This plan, "founded on the basis of the friendly societies," would create a single compulsory mutual society for the entire country with the aim of promoting individual and collective security.

...[I]t seems but just that every man should be obliged, out of his present Abundance, to take so easy a Method of securing to himself such a Subsistence, as, should his Necessities ever call for it, would insure the Public from his ever becoming a Burthen to it; and if they should not call for it, the gratuitous Contributions arising from the Subscription of all those, who are, by their Degree in Life, placed above the Receipt of any Return, would come happily in Aid of the poorer Contributors, and supply them with Means of a better Allowance under all their Wants.¹³¹

Like reformed friendlies, Acland's national scheme would reflect both personal responsibility and the limited altruism of the shared risk pool. In addition, its large public scale would allow for "a national Security for its accumulating Capital," compound interest on its funds, and "what is of no small Weight, a very considerable Resource, to eke out the Poor Man's Subscriptions, in the gratuitous Contributions of the Richer Subscribers."¹³² Just as voluntary mutual insurance was less risky and more collectively beneficial than private savings, compulsory social insurance would on this account be more secure, more encompassing, and even more altruistic than the friendlies.

Meanwhile, in France, the 1780s and 1790s also saw a variety of national insurance proposals. One of the first was that of Andre Jean de Larocque, who in 1785 proposed a system of general savings accounts that would invest funds generated by regular workers' contributions and

¹³¹ John Acland, *A Plan for Rendering the Poor Independent on Public Contribution; Founded on the Basis of the Friendly Societies, Commonly Called Clubs* (London: 1789), 11.

¹³² *Ibid.*, 28.

return the proceeds as annuities.¹³³ Shortly thereafter, Marie-Jean-Antoine-Nicolas Caritat de Condorcet proposed probabilistic social insurance along similar lines.¹³⁴ In his late and unfinished *The Sketch*, published in 1795, Condorcet cites Richard Price, along with fellow liberals Jacques Turgot and Joseph Priestly, as the “first and most brilliant apostles” of the new doctrine of “the indefinite perfectibility of the human race,” which would eventually change social relations by putting an end to the arbitrary dependence and persistent inequality caused by ignorance and prejudice.¹³⁵ Condorcet believed that the probability calculus would play a central role in this process, providing the tools to illuminate and conquer the workings of nature, to improve individual reason, to rationalize political economy, and perhaps above all to promote equality of conditions and the “true perfection of man.”¹³⁶

Condorcet introduces his proposal by recognizing the insecurity of workers’ livelihoods and arguing that the dependence and suffering they experience whenever they lose their work could be “in great part eradicated” through publicly orchestrated mutual insurance. Such methods had “already been successful” thanks to “the application of the [probability] calculus to the probabilities of life and the investment of money.” But they had not yet “been applied in a sufficiently comprehensive and exhaustive fashion to render them really useful...to society as a whole.”¹³⁷

¹³³ Keith Michael Baker, *Condorcet: From Natural Philosophy to Social Mathematics* (Chicago and London: University of Chicago Press, 1975), 280-81.

¹³⁴ In 1790, Condorcet had also advocated for the establishment of accumulation accounts to facilitate government borrowing and promote economic activity by reducing the incentive for individuals to hoard their savings. *Ibid.*, 281.

¹³⁵ Jean-Antoine-Nicolas Caritat, Marquis de Condorcet, “The Sketch,” in *Condorcet: Political Writings*, ed. Steven Lukes and Nadia Urbinati (Cambridge: Cambridge University Press, 2012), 102.

¹³⁶ *Ibid.*, 125.

¹³⁷ *Ibid.*, 131. Condorcet almost certainly had in mind life insurance mathematics in citing the demonstrated success of probability theory as the inspiration for his own plans. He was well aware of the work of the Equitable, having cited it in his works on probability theory, and had written an extensive essay on the mathematics of calculating insurance rates based on probabilities. See Condorcet, “Assurances (maritimes),” in *Arithmétique politique: Textes rares ou inédits (1767-1789)*, ed. Bernard Bru and Pierre Crépel (Paris: Institut National D’études Démographiques, Presses Universitaires de France, 1994), 485-494.

The three schemes that Condorcet goes on to propose closely resemble actuarial insurance both in their spirit and in their workings. The first is a kind social security, which would guarantee the elderly “a means of livelihood produced partly by their own savings and partly by the savings of others who make the same outlay, but who die before they need to reap the reward.” The second would secure “for widows and orphans an income which is the same and costs the same for those families which suffer an early loss and for those which suffer it later.” And the third would provide children with “the capital necessary for the full use of their labor,” which increases “at the expense of those whom premature death prevents from reaching this age.”¹³⁸ Technically, these schemes are designed to function like annuities, in that they rely on the fact that some participants will die before collecting their benefits and therefore subsidize those who end up living longer than average. Yet they also have all of the major characteristics of insurance: a payment calculated as a function of at least one of the insured’s characteristics, designed to compensate for a prospective misfortune or loss, and made to a party whose sole role in the agreement is to shoulder the burden of that loss.

Condorcet justifies the need for such programs on the ground that those who depend on their own labor to survive are far more vulnerable to the destructive effects of chance than those whose income derives from other sources and “whose resources are not at all subject to the same risks.”¹³⁹ The inequality and dependence that result from this state of affairs are therefore unnecessary and undeserved, a proper target for ameliorative efforts. While it would be “foolish and dangerous to wish to eradicate” all inequality, since some is the result of “natural and necessary causes,” social insurance promises to eliminate arbitrary inequality and promote a kind of universal self-sufficiency, such that everyone is able to “manage his household, administer his affairs, and

¹³⁸ Condorcet, “The Sketch,” 131

¹³⁹ Ibid.

employ his labor and his faculties in freedom.”¹⁴⁰ Whatever inequality does survive will be in everyone’s interests, promoting “the progress of civilization...without entailing either poverty, humiliation or dependence.”¹⁴¹

On Condorcet’s account, then, the purpose of social insurance is to prevent severe inequality and its concomitant dependence. Although his argument is in important respects liberal, he contemplates both compulsory and voluntary schemes without any apparent sense that the first might violate individual freedom. The liberalism of this account therefore hinges not on the voluntary choice to insure, but rather on the assumption that insurance is a rational decision (whether or not it is in fact chosen) and on the claim that it provides a kind of freedom one could not secure without it. As we will see in greater detail below, the liberalism of Condorcet’s account also rests implicitly on a conception of risk as calculable for individual events, and of insurance as a bilateral contract between the insured and his insurer. While requiring collective participation, then, insurance remains on this view a tool for securing one’s own welfare through calculated provision against the vicissitudes of chance.¹⁴²

Another important proposal for welfare policy made at the same time as Condorcet’s came from Thomas Paine’s *Agrarian Justice*, also published in 1795. In it, Paine offered an attempt to alleviate poverty through a system of universal benefit payments.¹⁴³ His idea was to create a national fund, supported by an inheritance tax, out of which all citizens would receive a lump sum at the age

¹⁴⁰ Ibid., 132.

¹⁴¹ Ibid., 133, 126.

¹⁴² On Condorcet’s economic liberalism, grounded in the natural rights of personal liberty and the secure enjoyment of property, see Baker, *Condorcet*, 218-19.

¹⁴³ Elizabeth Anderson has called this “the first realistic plan to abolish poverty on a nationwide scale” and a precursor of modern social insurance. Elizabeth Anderson, “Thomas Paine’s ‘Agrarian Justice’ and the Origins of Social Insurance,” in *Ten Neglected Classics of Philosophy*, ed. Eric Schliesser (Oxford Scholarship Online, 2016), DOI: 10.1093/acprof:oso/9780199928903.003.0003.

of 21 and everyone over the age of 50, along with anyone who is “blind and lame,” would receive annual fixed payments. Paine described these benefits as compensation for the loss of all men’s “natural inheritance” following “the introduction of the system of landed property,” which has “has absorbed the property of all those whom it dispossessed, without providing, as ought to have been done, an indemnification for that loss.”¹⁴⁴ Paine’s radical critique of the agrarian property distribution led him to view these payments as “not charity but a right, not bounty but justice.”¹⁴⁵ Their purpose was to offer a form of redress for the suffering caused by what was essentially the theft of common property through private appropriation.

In trying to protect workers from the risk of destitution and compensating all citizens for arbitrary inequalities, Paine might also seem to be proposing an early version of social insurance. Yet there are also important differences between his scheme and anything that we can readily recognize as insurance. Whereas in principle insurance offers protection against a future loss, Paine’s program sets out to redress a prior injustice. Whereas insurance calibrates (however roughly) the insured’s payments to the total benefit he can expect to receive, Paine’s scheme draws its funds from those who no longer have anything at stake and redistributes them evenly to those who have contributed nothing. Whereas insurance equalizes individuals insofar as they share certain characteristics, Paine assumes a kind of primordial equality—expressed in the assumption of original common ownership—that has nothing to do with empirical likeness. And whereas insurance operates in the realm of contract, Paine tends to discuss his proposal in the language of tort.

In addition, Paine’s critique of property ownership distinguishes his vision rather sharply from that of Condorcet and others who saw social insurance as aligned with an essentially bourgeois

¹⁴⁴ Thomas Paine, *Agrarian Justice* (Paris: 1797), in *The Writings of Thomas Paine, Vol. III (1791-1804)*, ed. Moncure Daniel Conway (New York: G. P. Putnam’s Sons, 1894), 330, <http://oll.libertyfund.org/titles/paine-the-writings-of-thomas-paine-vol-iii-1791-1804>.

¹⁴⁵ *Ibid.*, 336.

ethic. Insurance in its traditional form assumes the legitimacy of private property by recognizing not only the good being insured, but also private ownership of the risk itself.¹⁴⁶ Paine's scheme thus locates arbitrary and compensable injustice much earlier in time than a mutual insurance society that compensates only losses anticipated and paid for by the individual. Insofar as it is linked, even partially, to an account of personal responsibility as providing for oneself against calamity, insurance invokes and reinforces those middle-class virtues whose exercise depends on private property.¹⁴⁷

That Paine did not propose a version of social insurance should not necessarily rule him out as a progenitor of the modern welfare state. It does, however, remove him from the line of thinking under consideration here, which aimed to realize a vision of distributive justice as forward-looking mutual support among probabilistic equals. Like the liberal political tradition out of which it grew, this vision set out to align individual interest with the common good through universal fear or concern for the future, combined with the industrious, orderly behavior to which it can, when properly channeled, give rise. A risk-sharing association governed by the impersonal laws of mathematics was both an encapsulation of this liberal vision and a tool for its realization.

II. The moral character of late-classical probability

One of the central arguments of this dissertation is that thinking about probability, and therefore about risk and insurance, has always been intimately bound up with moral and political aims. In the case of classical probability, as we saw in the last chapter, it was the legal tradition of

¹⁴⁶ See Jonathan Levy, *Freaks of Fortune: The Emerging World of Capitalism and Risk in America* (Cambridge, MA: Harvard University Press, 2014).

¹⁴⁷ It is revealing in this respect that Condorcet refers three times to Locke as the foundational authority of modern philosophy: first in revealing the "universal moral law" by which "efforts made by each individual on his own behalf minister to the welfare of all"; second in his "precise analysis of ideas," which cut through the "chaos of incomplete, incoherent, and indeterminate notions which chance presents to us at hazard"; and finally in allowing for the application of reason to "moral science, to politics and to social economy." Condorcet, "The Sketch," 93, 95, and 96.

aleatory contracts and the need for an accounting of just proceeds that inspired mathematical probability theory and gave rise to the concept of expectation. In the tradition that succeeded it, I will now argue, it was the normative vision of social mathematics and its particular brand of enlightened liberalism that motivated important parts of the probability calculus.

The tradition in question, which I refer to as late-classical or *a posteriori* probability, employed elements of the classical method to calculate predictive likelihoods on the basis of prior empirical observation. It continued to use the device of equiprobable cases, and with it to calculate likelihoods for single events, while at the same time regarding probabilities as characteristics of large statistical groups. As a result of this duality, thinkers in this tradition were able to maintain a basically liberal rationale for insurance while at the same time reimagining it as a tool for collective political ends. The resulting account presented insurance as a prudent individual choice, entered into for personal material benefit, while also promoting the idea that to insure is social obligation. In the next section, I will analyze in greater detail the arguments that these thinkers produced on behalf of social insurance as both an individual and a political concern. First, however, I will consider how the practical aims of late-classical probabilists inflected their approach to calculating risk, and how their interpretation of probability therefore both mirrored and furthered their political project.

A. Inverse probability

The normative concerns animating late-classical probability can be clearly appreciated by examining its most significant technical innovation: the theory of inverse probability or the probability of causes. This was an approach, pioneered by Pierre-Simone Laplace and Thomas Bayes, for calculating the likelihood of an unknown composition or state of the world given the observed outcomes of events. It is a form of induction—that is, of inferring from a number of

cases of which something is true that the same is true of the entire class—and the foundation of what one would today call statistical induction. In the hands of its pioneers and early adopters, particularly Condorcet and Laplace, it seemed to hold the key to a range of potential applications of probability to practical affairs, not least of all insurance.

In a short manuscript on the history of probability, Condorcet credits Jacob Bernoulli for taking the first step in redirecting the discipline toward its new, empirical focus.¹⁴⁸ Bernoulli had found that for trials satisfying the axioms of the binomial distribution—that is, where a series of independent outcomes can each yield only a positive or negative result—the observed frequency of a given outcome converges to the true probability as the number of observations increases to infinity. Bernoulli concluded from this that where the number of trials is infinitely large, the observed frequency is as good an estimate of the true probability as the *a priori* value.

This theorem, however, was only suited for estimating the degree of likelihood that, given a number of trials, the observed frequency of events would differ within certain limits from the true one.¹⁴⁹ Abraham De Moivre subsequently improved on Bernoulli's approach, finding that where one can assume an underlying frequency, "the ratio of happenings will continually approach to that [value], as the experiments or observations are multiplied."¹⁵⁰ Yet his demonstration was logical rather than mathematical. Thus while earlier treatments had suggested the possibility of a theory of statistical induction, they had not yet offered the mathematical tools for putting it into practice.

¹⁴⁸ Condorcet, "Petit ouvrage 'plus métaphysique que mathématique' sur le calcul des probabilités," in *Arithmétique politique*, 282-294.

¹⁴⁹ Bernoulli did not provide a way of using his theorem to find an interval for the true probability based on an observed frequency. Nevertheless, he tried to extend his theory well beyond these somewhat narrow confines. For one thing, he proposed relative frequency as an estimator of true probability in other applications, even though *a posteriori* calculation was justified only for trials with a finite number of equally likely outcomes. See Hald, *A History of Probability and Statistics*, 258-59, 263.

¹⁵⁰ See Abraham de Moivre, *Doctrine of Chances*, 3rd ed. (London: 1756), 243-254 and Karl Pearson, "James Bernoulli's Law," *Biometrika* 17, no. 3/4 (1925): 205-6.

Those tools came from the independent work of Laplace and Bayes. Though Bayes's contribution came first, Laplace's articulation was originally better known and more influential (and will therefore be the focus of my discussion here).¹⁵¹ In a 1774 memoir on the subject, Laplace analogizes real-world events to black and white balls in an unknown proportion selected at random from an urn.¹⁵² Assuming that the balls are replaced after each draw, Laplace first sets out to determine the probability that the true ratio of balls in the urn is a given number x , in light of the observed outcomes. Thus if one had witnessed that out of 100 drawings, 30 balls were black and 70 white, Laplace shows how to determine from this finding the likelihood that the true ratio of balls is 3:7, or any other ratio one might choose. If x is the (unknown) ratio of balls, then the probability of drawing p white and q black tickets is $x^p(1-x)^q$. To find the probability that x is indeed the true ratio, one simply divides this figure—the likelihood of witnessing the observed outcome given an underlying ratio of x —by the likelihood of that observed outcome given all possible underlying ratios.

So far, Laplace's result is the equivalent of what is now known as Bayes's theorem, with the assumption that all ratios or "causes" are *a priori* equally likely. Laplace has shown how, given very restrictive assumptions, one can calculate the probability that some observed phenomenon (say, a record of 51 percent male births in a county) results from a particular underlying state of affairs (for example, an equal likelihood of male and female infants in nature as a whole).

¹⁵¹ Bayes's *Essay towards Solving a Problem in the Doctrine of Chances* was written in the 1740s but published posthumously, in 1764, by none other than Richard Price. According to Stephen Stigler, debates about inverse probability through the 19th century centered on Laplace's work and not Bayes's, whose contributions languished in relative obscurity. Major contributors to the field such as Boole, Venn, Jevons, and Edgeworth all took pains to discuss Laplace, but none made substantive references to Bayes. Stephen Stigler, "Thomas Bayes's Bayesian Inference," *Journal of the Royal Statistical Society* 143 (1982): 255. Laplace's memoir also had a profound effect on Condorcet, inspiring him to regard probabilities as the foundation for a new mathematical science of conduct. See Baker, *Condorcet*, 171.

¹⁵² Laplace, "Memoire on the Probability of the Causes of Events," trans. Stephen M. Stigler, *Statistical Science* 1, no. 3 (1986): 364-78. The following account focuses on Laplace's discussion on 364-70.

This finding is not predictive, however, so Laplace pressed on, seeking to derive the probability of drawing another white ball given prior draws of p white and q black balls. Again assuming x to represent the unknown ratio of black to white balls, the probability of drawing an additional white ball from an urn with that ratio can be expressed by $x^{p+1}(1-x)^q$. Multiplying this expression by the probability of x , as defined above, gives the probability of drawing another white ball from the urn in question (since conditional probabilities are always calculated by multiplying the probability of the desired outcome by the probability of the prior condition). Finally, taking the integral of the numerator (the expression representing the probability of another white draw given x) for all values of x between 0 and 1 generates the total probability of drawing another white ball, accounting as it does for all possible values of x . One can generalize this result even further, Laplace shows, to the probability of drawing m white and n black balls from the urn given p and q prior draws. This formula, which became known as the Rule of Succession, allows the practitioner to calculate the likelihood of any desired future outcome given certain empirical inputs and the assumption that all possible prior states of the world are equally likely to have “caused” the results.¹⁵³

Nor does Laplace stop here. He goes on to simplify his results even further, showing through a series of approximations that where p and q are very large and m and n very small, one can conclude that the true underlying ratio of balls is equal to $p:q$. In other words, under certain conditions one need no longer assume that the underlying ratio of balls is unknown, but can reasonably operate as though the underlying or real-world frequency of events equals the observed one. In affirming the identity between “real” probabilities and long-term observed ones, Laplace

¹⁵³ As subsequent commentators pointed out, and as Condorcet recognized as well, Laplace’s formula considers only the probability for a drawing of m white and n black balls from the urn in that particular order. To account for the probability without regard to the order in which the $m + n$ balls are drawn, one must add another term, the binomial coefficient, before the expression that Laplace derives here. See Andrew I. Dale, *A History of Inverse Probability from Thomas Bayes to Karl Pearson*, 2nd ed. (New York: Springer-Verlag, 1999), 172; Isaac Todhunter, *A History of the Mathematical Theory of Probability, from the time of Pascal to that of Laplace* (Cambridge and London: Macmillan and Co., 1865), 384; and Condorcet, *Memoire sur le Calcul des Probabilités*, in *Arithmétique politique*, 420.

both reflected and encouraged the movement away from the classical understanding of probability and its interpretations of expectation and equiprobability.¹⁵⁴ At the same time, we will now see, he continued to maintain that the calculus could be a guide to individual decisions, and in so doing preserved its liberal moral character despite encouraging its aggregative, statistical turn.

B. The principle of indifference

As we just saw, the attempt to determine unknown and predictive frequencies on the basis of observed evidence—the central project of inverse probability—requires assigning prior probabilities to the possible causes in question. In both Laplace’s and Bayes’s analyses, this took the form of assuming what is now known as a uniform prior distribution, meaning that every possible cause or proportion is considered at the outset as equally likely to be true one. Returning to the famous urn example, if one imagines a different urn for every possible proportion of balls, Laplace’s formula considers it just as likely that the observed selection came from an urn containing exclusively black balls as from an urn containing only white ones—and the same for all possible ratios in between.

Laplace never offered an explicit justification for this assumption, though he continued to employ it in many of his calculations.¹⁵⁵ The only plausible explanation to be found anywhere in his work stems from what has come to be known as the principle of non-sufficient reason or

¹⁵⁴ Daston similarly characterizes Laplace as a transitional figure in the history of probability theory, concluding that his work, “particularly his enthusiasm for statistics, marked the beginning of the end for the classical interpretation that had dominated eighteenth-century probability theory.” Daston, *Classical Probability*, 284. According to Hacking, however, the so-called classical theory was in fact a set of competing theories that shared a single classical feature, namely the definition of probability in terms of a fundamental probability set, or a set of disjointed, equally probable alternatives, which Laplace maintained. Ian Hacking, “Jacques Bernoulli’s Art of Conjecturing,” *British Journal for the Philosophy of Science* 22, no. 3 (1971): 210.

¹⁵⁵ Unlike Bayes, who had provided a fairly extensive argument for his assumption of equal prior probabilities, Laplace treats it as an intuitive axiom. This is one point that has prompted scholars to disassociate their two accounts. See Stephen M. Stigler, “Introduction,” in “Laplace’s 1774 Memoir on Inverse Probability,” *Statistical Science* 1, no. 3 (1986): 359 and “Thomas Bayes’s Bayesian Inference,” *Journal of the Royal Statistical Society* 145, no. 2 (1982): 250-258.

indifference.¹⁵⁶ Although not an argument for the equal likelihood of causes per se, this is the general principle that, absent knowledge to the contrary, one ought to assume that all possible events are equally likely to occur. “When we wish to make use of this theory,” Laplace wrote in his 1774 memoir, “we regard two events as equally probable when we see no reason that makes one more probable than the other.”¹⁵⁷

Such a principle had no reason to exist in the realm of *a priori* probability, where all likelihoods are known in advance.¹⁵⁸ It is only with the transition to empirical probability, and in particular the effort to deduce underlying and predictive likelihoods from observed ones, that a mathematical expression for the calculator’s ignorance became necessary. Laplace, for his part, seems to have felt that the principle reflected the skeptical, open-ended nature of the scientific enterprise, and thus strengthened the link between inverse probability and the natural sciences.¹⁵⁹ Indeed, it appears to have been a direct outgrowth of epistemological currents of the time,

¹⁵⁶ Long known as the principle of non-sufficient reason, it was Keynes who first designated it as the principle of indifference, a term I also prefer for its economy of expression. See John Maynard Keynes, *A Treatise on Probability* (Lexington, KY: Wildside Press, 2013): 45.

¹⁵⁷ Laplace, “Memoir on the Probability of the Causes of Events,” 378.

¹⁵⁸ According to Anders Hald, Jacob Bernoulli offered an early version of the principal of indifference, arguing that causes must be assumed equally possible to avoid the need for discretion in determining the true probability of events. See Hald, *A History of Probability and Statistics*, 249-252. Whether Bernoulli in fact intended to describe a situation in which one is truly indifferent among possible alternatives because ignorant of their respective likelihoods, or whether on the contrary he considered the alternatives to be equally probable in an objective, physical sense, is taken up in Hacking, “Jacques Bernoulli’s Art of Conjecturing,” 219.

¹⁵⁹ This is Daston’s explanation, although her account of why Laplace adopted the principle is unconvincing. First, she asserts that Laplace offered a “restatement” of Bayes’ scholium concerning equal prior probabilities, but offers no evidence that Laplace regarded his own treatment of the principle as such. Indeed, the work to which she cites in support of this claim was composed after Laplace’s original work on inverse probabilities, in which he was apparently unaware of Bayes’s work but nevertheless made the same assumption that all possible ratios of x should be considered equally likely. Second, Daston’s claim that Laplace’s “interpretation of Bayes’ theorem as a mathematicized scientific method may have smoothed the way for his sweeping assumption of equal prior probabilities” makes little sense on the view that Laplace was unaware of Bayes’s work before composing his seminal paper in 1774. Even the more plausible argument (not made explicitly by Daston) that Laplace reinterpreted his own assumption in light of Bayes’s work is undermined by the fact that Laplace persisted throughout his life in defending his principle as a consequence of *ignorance*, while Bayes’s argument for the assumption of a uniform prior distribution did not. See Daston, *Classical Probability*, 274, and Stigler, “Thomas Bayes’s Bayesian Inference.”

combined with the ineradicable need of probability theory for a common denominator of equally likely alternatives.

By this time, probability had long been linked to an associationist epistemology, which presented knowledge of the natural world as a product of connections between experiential facts.¹⁶⁰ While such knowledge could never attain demonstrative certainty, the more one observed a connection between two events, the more likely one could consider it to be that they would accompany each other in the future. Laplace and his contemporaries understood mathematical probability as the scientific expression of this reasoning process—and thus, in Laplace’s formulation, as relative “in part to this ignorance” (i.e., of true causes), “in part to our knowledge” (as derived from experience).¹⁶¹ While conceding the human need to make inferences from experience, this approach remains ultimately skeptical about the past’s reliability as a guide to the future.

In this light, then, the principle of indifference can be seen as above all a practical tool, the necessary precondition for scientific reasoning rather than its conclusion. By giving numerical expression to the calculator’s initial ignorance, it allows him to generate inferences that would be unavailable without it. As such, it also reflected a new interpretation of the equipossibility model at the foundation of mathematical probability. Originally understood as an aleatory condition, a representation of the actual chances of an event, in Laplace and his successors it became an epistemic one, a representation of ignorance regarding those chances and indeed of nature’s inner workings more broadly. Calculations based on this epistemic assumption would ultimately

¹⁶⁰ On probability in Locke, for example, see Douglas John Casson, *Liberating Judgment: Fanatics, Skeptics, and John Locke’s Politics of Probability* (Princeton: Princeton University Press, 2011), 103-14. Even Hume, who denied the rationality of inductive inference, took probabilistic thinking based on repeated observation to be the standard of reasonable judgment. See David Hume, *An Enquiry Concerning Human Understanding*, Sect. VI, in *Essays and Treatises on Several Subjects* (London: 1777); Daston, *Classical Probability*, 193-96; and Casson, *Liberating Judgment*, 154.

¹⁶¹ Laplace, *Philosophical Essay on Probabilities*, 3. Hence also Laplace’s assertion that mathematical probability is “basically only common sense reduced to a calculus. It makes one estimate accurately what right-minded people feel by a sort of instinct, often without being able to give a reason for it.” *Ibid.*, 124.

correspond with actual probabilities in the world, harmonizing individual judgment with scientific observation.

Indeed, in the context of late-classical probability, the principle of indifference effectively served as the bridge between probability's epistemic and aleatory interpretations. Specifically, it allowed for the continued use of probabilities to guide individual judgment and choice despite the newly statistical cast of the discipline. It did this in two ways: first, by providing the basis on which the Rule of Succession's predictive probabilities rested; and second, by stating that for any eventuality about which one is uncertain, one can reasonably assign equal probabilities to its possible outcomes.¹⁶² Thanks to these conclusions, individuals could continue to use mathematical likelihoods to guide their own decisions even as probability calculations focused increasingly on long-run frequencies.

The principle of indifference can therefore be seen to encapsulate the moral and political aims of thinkers such as Condorcet, Laplace, and other so-called social mathematicians who envisioned that universal enlightenment would lead to a rational social order in which individual freedom aligns with the common good.¹⁶³ As Condorcet put it, "What are we to expect from the perfection of laws and public institutions...but the reconciliation, the identification of the interests of each with the interests of all?"¹⁶⁴ Social insurance was just one expression of this hope; the marquis also sought to use probabilities to rationalize judicial decisions and voting rules. Applying the calculus to facts that we have not seen ourselves "teaches us to find out and to measure the true strength of our reasons for believing" them, shedding "the light of reason on areas which have too

¹⁶² In his 1774 essay, Laplace had emphasized that to calculate the probability of more than one event's occurring, empirical rather than *a priori* probability values must be used. Yet for the single events, one could still make use of an *a priori* value—specifically, the equal probabilities enabled by the principle of indifference. Laplace, "Memoire on the Probability of the Causes of Events," 378.

¹⁶³ See also the discussion in Baker, *Condorcet*, 225-244.

¹⁶⁴ Condorcet, "The Sketch," 140.

long been abandoned to the seductive influence of imagination, self-interest and passion.”¹⁶⁵

Mathematics could vanquish that “empire imposed by...passions over truth, by active ignorance over light,” replacing mere opinions with the universally compelling logic of calculation.¹⁶⁶

Ultimately, however, the presumed harmony between the epistemic and aleatory aspects of probability was just that—presumed. It is not always the case that a mathematical likelihood derived from empirical observation will provide a sound guide to individual choice; in fact, such likelihoods may well lead individuals astray. This problem is not unique to *a posteriori* likelihoods, but it was heightened by late-classical probability’s emphasis on large-scale observation as the basis for its calculations. As we will now see, the technical bridge provided by the principle of indifference found a parallel in the device of moral expectation, and by extension the idea of a duty to insure.

C. Mathematical and moral expectation

The possibility of a divergence between the dictates of prudence or economic reason and equity as defined by probability calculations has been a subject of concern for almost as long as the existence of probability itself. It was first raised in the context of the St. Petersburg paradox, in a 1713 letter from Nicholas Bernoulli to Pierre Montmort. The paradox involved two players in a coin-toss game, the rules of which are that one player gives the other two cents if heads turns up on the first toss, four if heads first turns up on the second toss, eight if on the third, and so on. If heads does not turn up until the n th toss, the first player owes the second 2^{n-1} cents. According to the standard method of calculating expectation, the second player’s expectation from this game is actually infinite, since it is possible (if very unlikely) that a fair coin will land heads an infinite

¹⁶⁵ Condorcet, “Le projet reformulé: Tableau général de la science qui a pour objet l’application du calcul aux sciences politiques et morales,” in *Mathématique et Société*, ed. Roshdi Rashed (Paris: Hermann Editeurs, 1974), 205.

¹⁶⁶ *Ibid.*, 197-98.

number of consecutive times. Yet we all know that no levelheaded person would pay an infinite sum of money for that player's stake in the game. There consequently appears to be a conflict between expectation as the determinant of fairness and the considerations that in fact move a prudent decision-maker.¹⁶⁷

In response to this difficulty, Nicholas's cousin Daniel Bernoulli proposed a distinction between mathematical expectation, with its goal of contractual equity, and moral expectation, which takes into account an outcome's anticipated utility to the actor. In the case of the St. Petersburg game, Bernoulli argued that the player's expected benefit from playing depends on his initial wealth, with the value being lower the smaller his initial fortune and increasing with diminishing margins as base income goes up. He thus explained the difference between the dictates of contractual equity and the empirical psychology of risk-taking by personalizing the value of the wager to the decision-maker himself.¹⁶⁸ He also showed that, assuming personal utility curves are always concave, games of chance judged fair by the standard of mathematical expectation alone in fact involve negative utility for both players.¹⁶⁹

Daniel Bernoulli's argument can be understood as an attempt to shift the focus of mathematical analysis from equity to utility or economics.¹⁷⁰ Unlike the original probability calculus, in which "no characteristic of the persons themselves ought to be taken into consideration," moral

¹⁶⁷ See also Daston, *Classical Probability*, 70.

¹⁶⁸ Although he acknowledged the difficulty of generalizing with regard to moral expectation, since "the utility of an item may change with circumstances," he devised his formula on the assumption that, in general, human beings appreciate less each additional unit of wealth the greater their initial fortune. Daniel Bernoulli, "Exposition of a New Theory on the Measurement of Risk," trans. Louise Sommer, *Econometrica* 22, no. 1 (1954): 24-25.

¹⁶⁹ *Ibid.*, 23-36.

¹⁷⁰ See Daston, *Classical Probability*, 70-76.

expectation accounts for “the particular circumstances of the person making the estimate.”¹⁷¹ The former is the province of the law, the latter of the individual.¹⁷²

Statistical life insurance, by contrast, implicitly shifts the focus in the other direction—to equity understood from the perspective of the insurance company or social planner. Among the first to recognize this was Jean de la Rond D’Alembert, a rare early philosophic skeptic about the ability of probability theory to model or guide prudential reasoning at its best. Analyzing the rules for calculating lifespans, D’Alembert pointed out that the same numerical average might result from a number of different life prospects. For example, in a group of 50 people who live a total of 100 years, the life expectancy of each one is two years. Yet it is quite a different matter to know with certainty that one will die in two years than to know that one has a 50 percent chance of living for four.¹⁷³ Expectation defined as the average lifespan fails to capture the difference.

D’Alembert’s critique of statistical averages thus returns to the perspectival dichotomy raised by Daniel Bernoulli. Such averages may be useful to an insurance company or social planner acting on behalf of the aggregate good, but they are far less useful to the individual acting on behalf of his own good. D’Alembert put the point particularly starkly in his discussion of the smallpox inoculation controversy raging at the time. Envisioning a scenario in which inoculation increases the life of four out of five people from 50 to 100 years but causes one of the five to die right away, D’Alembert explains that no individual would freely choose inoculation even though the state would

¹⁷¹ D. Bernoulli, “New Theory on the Measurement of Risk,” 24.

¹⁷² Ibid.

¹⁷³ The alternative formula for life expectancy—the number of years at the end of which half of those currently living will die—faces no less significant difficulties on D’Alembert’s view. Here, the average value is the same whether m living people die seriatim until a half are alive at the end of p years, or all live p years at the end of which half suddenly die. Here again, in one case there is a certainty of living p years (and a possibility of living longer) while in the other there is only a chance or expectation of living that long. Jean de la Rond D’Alembert, “Sur la durée de la vie,” in *Opuscules mathématiques*, vol. IV (Paris: 1768), 97, reprinted in *Arithmétique Politique dans la France du XVIIIe siècle*, ed. Thierry Martin (Paris: Institut National D’Études Démographiques, 2003), 245.

prefer it. While the latter takes the perspective of the aggregator, “considering all citizens indifferently,” for the individual “the interest of his particular conservation is foremost,” and an increase of the average lifespan is not in itself determinative.¹⁷⁴

The conventional approach to calculating expectation goes astray, on D’Alembert’s view, because it does not properly weigh outcome values in accordance with their probabilities. Multiplying the two together assumes they are equally significant, but prudence requires considering the likelihood of success or failure *in relation to* the potential reward or loss, a relation that D’Alembert calls the risk. Where the probability of gain is extremely small, a reasonable person will not pay any significant amount for the wager, even if the potential gain is enormous. Similarly, in the case of inoculation, what determines the reasonableness of the choice is “uniquely the relationship between the risk on one side, and on the other the augmentation of the average life, or above all the advantage that this augmentation will procure relative to the time and the age at which one will enjoy it. The difficulty is to determine this relationship.”¹⁷⁵

The decision of whether to insure one’s life is far less stark than D’Alembert’s inoculation example, and the expectations determined from mortality tables do convey more information than the average lifespans he describes. Yet his concerns about the limitations of aggregation and abstraction apply to statistical life insurance as well. Distributive justice on the probabilistic account hinges on the parties’ abstract potential to “win” or “lose” the common resource pool. Where probabilities were based on coin tosses and similar examples, it was easy to imagine that these potentials could be calculated, and the parties equalized, with mathematical precision. As empirically derived probabilities replaced *a priori* ones, however, it became more difficult to understand

¹⁷⁴ D’Alembert, “Sur l’application du Calcul des Probabilités à l’inoculation de la petite Vérole,” in *Opuscules Mathématique*, vol. II, bk. 11 (Paris: 1761), 37-38.

¹⁷⁵ *Ibid.*, 36.

expectation as the precise expression of a party's payoff. For the life insurance company, like the state, the benefits of treating individuals as at a certain point interchangeable are clear. For the individual, by contrast, there is quite a bit of relevant information that the insurance company's aggregative stance fails to capture.¹⁷⁶

In response to this problem, thinkers in the tradition of late-classical probability consistently turned to moral expectation to justify insurance as an individually rational choice. Thanks to this rule, Laplace explained, one can determine “how great the sacrifice that ought to be made to the insurance company should be” above and beyond one's mathematical expectation “if a moral advantage is always obtained.”¹⁷⁷ Similarly, physicist and mathematician Jean-Baptiste Joseph Fourier argued that to determine the fair price of an insurance premium one must find, “by an exact calculation, what must be the contributive portion of each associate” as a function of both the average group risk and each individual's personal utility.¹⁷⁸ Moral expectation, as we will see in greater detail in the next section, could also support the claim that there is a duty to insure, both to oneself (given the undisputed benefits of security) and to the group on which one's security depends.

We have thus seen that late-classical probabilists set out to reconcile the attractiveness of insurance as an individual choice with its promise of stability for the group. They did so by introducing within the calculus the means to align personal risk assessments with objective frequencies and their collectively beneficial uses. Although their solutions are certainly vulnerable to

¹⁷⁶ Daston points out that while in its early years the Equitable sometimes varied its premiums on the basis of in-person interviews with potential policyholders, its successors abandoned the practice as their faith in the predictive power of long-run averages grew stronger. See Daston, *Classical Probability*, 180-183.

¹⁷⁷ Laplace, *Philosophical Essay on Probabilities*, 88-9.

¹⁷⁸ J. B. J. Fourier, “Extrait d'un Mémoire sur la théorie analytique des assurances,” *Annales de Chimie et de Physique*, vol. 10 (1819), 177-78, 180, <http://gallica.bnf.fr/ark:/12148/bpt6k6570892b/f183.item.r=.zoom>. On the connection with Quetelet, see Stephen E. Feinberg, “A Brief History of Statistics in Three and One Half Chapters: A Review Essay,” *Statistical Science* 7, no. 2 (1992): 215.

critique, as later writers would forcefully point out, they allowed these thinkers to retain a basically liberal orientation while envisioning a kind of mutual dependence or reciprocity between the individual and the social whole. This vision was in turn central to their case for social insurance, to which we will now turn.

III. Early social insurance in theory and practice

As mathematical probability became increasingly statistical toward the end of the 18th century, a new account of insurance began to appear in probabilistic writings, and it was this account that inspired both the friendly society reformers and the social insurance proposals of the time. Rather than simply a bilateral transfer of risk, insurance could now also be seen as a means of spreading or pooling risks across a large group of people.¹⁷⁹ According to this view, as the insurer takes on more and more risks of a similar nature, his side of the transaction looks less and less like a wager, while the insured's payment decreases to reach his own mathematical expectation. According to the theory, the virtue of spreading or pooling risks in this way lies in its ability to align what the insurance association needs to collect from each member with the amount each should pay according to his personal risk assessment. At the limit, insurance ceases to resemble a gamble or even a commercial venture, and is simply a prudent and fair arrangement for the mutual sharing of burdens.

In reality, of course, such perfect alignment may not be attainable, even among non-commercial insurers. This happens when a risk pool is relatively small, increasing the chances that its members' claims will exceed the projected average, or when members are grouped imprecisely, meaning that the less risky end up subsidizing some of their peers. The challenge for probabilistic

¹⁷⁹ These terms reflect two sides of the same process: While the insured pool their risks together, the insurer spreads his risk over the expanded pool.

justice on a statistical view of risk is that the latter condition is almost invariably met. Likelihoods calculated from statistical data invariably ignore some features that meaningfully distinguish individuals from one another. Probabilistic insurance in general, and social insurance in particular, can therefore be seen as a kind of microcosm for thinking about political life more broadly. If there is a conflict between the strict justice of individual rewards and the security of the collective, which of the two should take priority? What does the individual owe the group—what if anything should she pay beyond her actuarial risk—and what does the group owe the individual in turn?

This problem saw a number of suggested resolutions in late-classical writings about insurance and the political proposals made in their wake. Some, as we have already seen, redefined prudence in terms of utility rather than mathematical expectation, justifying any additional price the insured might pay in terms of the psychological benefit of insurance to him personally. Others redefined justice in terms other than strict actuarial equity, prioritizing the security of the group over precision in calculating individual entitlements. Both sets of arguments together loosely constitute what I will call the liberal-statist account of social insurance that became prominent by the middle of the 19th century, particularly in France, and which inspired some of Western Europe's earliest welfare laws.

A. Between individual choice and social obligation

One of the first articulations of the risk-spreading rationale for insurance came from Condorcet. In an early article on the pricing of maritime insurance contracts, he proposed finding the fair price of an insurance policy by determining the long-run average amount that a merchant could pay while still profiting in his business, and the long-run average amount that an insurer had to earn for his enterprise to be sustainable. The fair price of the insurance in turn derives from “a

certain concurrence established” between these two amounts.¹⁸⁰ This use of long-run probabilities reduces the price of the policy by making the insurer’s outcome more certain. The risk of the insurers thus “spreading itself over a number of objects much larger than that of the insured, they can, in conserving a very large probability of gain, content themselves with a much smaller profit.”¹⁸¹

With this account, Condorcet had reimagined probabilistic justice as a limit problem, in which individual events yield disparate results but any discrepancy between the two sides decreases as the number of trials becomes infinitely large.¹⁸² Similarly, in one of his contributions to the *Encyclopédie Méthodique*, he defined the equality of two players in a game of chance as the equality of their long-run expectations, which can only be attained when the game is repeated a large number of times.¹⁸³ The implication is that no game can be fair when played only once; the only “safe” choice is to spread one’s bets over a large number of identical trials.

Two decades later, Laplace made a similar case for risk spreading in his own discussion of insurance. Mutual insurance, he argued, is the equivalent of sending small amounts of money on several ships rather than sending the entire sum on only one. “If two people, each having the same sum on two different ships sailing from the same port for the same destination, agree to distribute equally between them all the money that will arrive,” each will “distribute the sum he expects equally

¹⁸⁰ Condorcet, “Assurances (maritimes),” in *Arithmétique politique*, 488. The editors of this volume suggest that the work was drafted around 1783.

¹⁸¹ Ibid.

¹⁸² See also Condorcet’s manuscript Bibliothèque de l’Institut MS. 875, ff. 110v-114, summarized in Daston, *Classical Probability*, 98-99.

¹⁸³ Condorcet, “Probabilité,” in *Arithmétique politique*, 501. See also his *Essai sur l’application de l’analyse à la probabilité des décisions rendus à la pluralité des voix* (Paris: 1785), lxxvj, https://openlibrary.org/books/OL24141574M/Essai_sur_l'application_de_l'analyse_à_la_probabilité_des_décisions_rendus_à_la_pluralité_des_voix.

over the two ships.” As the number of associates increases, each one’s doubt about the possible loss decreases and his personal utility coincides “with the mathematical advantage, its natural limit.”¹⁸⁴

The mechanism of risk spreading remained central to the justification for insurance through the first decades of the 19th century. Silvestre Francois Lacroix, popularizing Laplace’s ideas in 1816, reiterated that “the larger the [mutual-aid] association” the more advantageous it will be, “in coming to form a fund that can ward off the irregularities that the succession of events presents in intervals...”¹⁸⁵ Augustus De Morgan, the prominent English mathematician who also served for a time as the Equitable’s actuary, relied on similar logic in 1838 in arguing against a proposed insurable interest law. Under such a law “it would be easy to gamble,” he wrote, since one can readily create a “bona fide insurable interest on a few lives,” but it would not be possible to spread one’s “bet” over a number of lives and thereby extend the transaction to the point where it ceases to be a wager. Only in allowing the “venturer to extend his traffic” will he “begin to *feel the average*” of his loss.¹⁸⁶

As both D’Alembert and Condorcet had recognized, however, this argument is much more naturally aligned with the perspective of the insurance company than with that of the individual insured.¹⁸⁷ Although in theory the individual’s premium will equal his expectation provided the pool is large enough, there is still the difficulty of identifying the average to which his experience is meant to refer. Condorcet, for example, presumes in his account of maritime insurance that each insurer and merchant will have sufficient experience that referring to an average personal profit makes

¹⁸⁴ Laplace, *Philosophical Essay on Probabilities*, 89. The same reasoning in turn recommends mutual over commercial insurance: While commercial insurers always charge more than mathematically necessary in order to make a profit, mutual insurers that spread their risk over a large pool—including over the nation as a whole—are in principle able to collect only their subscribers’ mathematical expectation.

¹⁸⁵ S. F. Lacroix, *Traite elementaire du calcul des probabilites* (Paris: 1816), 214.

¹⁸⁶ Augustus De Morgan, *An Essay on Probabilities* (New York: Arno Press, 1981), 103.

¹⁸⁷ Statistical probabilities, wrote Condorcet, are relevant “for states, whose interest embraces a large mass of men.” See Condorcet, “Tables de Mortalités,” in *Arithmétique politique*, 86.

sense, or that he can somehow incorporate the experience of similar insurers or merchants in making his decisions. Yet neither presumption is entirely sound, and Condorcet never attempts to justify them. The difficulty becomes even more pronounced in the case of insurance for death and other singular eventualities, for which no individual could possibly have enough experience to generate a long-run average of his own.¹⁸⁸

Writers in the late-classical tradition tended to follow Laplace in using moral expectation as a blanket solution to this problem. The psychological value of security, enabled by risk pooling, would resolve any discrepancy between the individual's own projected experience and the premium calculated for a group. This added value also distinguishes insurance from gambling, which merely exposes both parties to loss and its accompanying pain. As Laplace put it, "insurance, in which one exchanges the uncertain for the certain, ought to be advantageous" even when a fair game of the same mathematical amount is not, since in the latter "the player exchanges a certain stake for an uncertain gain."¹⁸⁹ Later, Belgian statistician Adolphe Quetelet, echoing his own teacher Fourier, reiterated that because in general the pain of loss outweighs the pleasure of gain, mutual insurance is morally advantageous while a gamble of the mathematically equivalent amount is not.¹⁹⁰ The "insured is generally guided by motives of prudence and economy; the gambler, by contrast, by improvidence and dissipation."¹⁹¹

¹⁸⁸ In his *Essai*, Condorcet gives a striking example of finding "the limit at which the danger of convicting an innocent and that of freeing someone culpable will be the same." But while in this case "society, if one will, would play an equal game, because it repeats an indefinite number of times," this will not be the same for the individual, who "can only play a number of times far too small for equality to be had for him." Condorcet, *Essai sur l'application*, lxxix.

¹⁸⁹ Laplace, *Philosophical Essay on Probabilities*, 88-9.

¹⁹⁰ In the same letter, Quetelet endorses a quotation from Buffon's *Essai d'arithmétique morale*: "The miser is like the mathematician: both estimate money by its numerical quantity; the sensible man considers neither its mass nor its number, he sees only the advantages he can derive from it; he reasons better than the miser and senses better than the mathematician." Quetelet, *Lettres sur la théorie des probabilités appliquée aux sciences morales et politiques* (Brussels: 1846), 49. See also Siméon-Denis Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités* (Paris: Bachelier, 1837), 8-9.

¹⁹¹ Quetelet, *Lettres*, 43.

These arguments from risk pooling and moral expectation could be seen to imply an obligation to insure, both to oneself and to the association, and indeed by the middle of the 19th century a number of thinkers had drawn precisely this conclusion.¹⁹² Not only is the failure to insure inexcusably self-destructive, since it could plunge oneself and one's family into misery far beyond the monetary loss incurred, but it also threatens the success of insurance for everyone else, exposing others to gratuitous risk. Statistical mutual insurance requires mass participation, since it realizes its optimal form only when a large number of similarly situated people join in together. It also depends on the capital created from their combined premiums, which are individually negligible but together have a power that exceeds the sum of its parts. "A centime is no more a capital than a wisp of iron is a lever. It is barely of value," wrote French journalist Edmond About in 1865. "You will find very few individuals who are sensitive to the loss or gain of a centime, because with an isolated centime one can do nothing. But he who by honest means would obtain this useless centime from his fellow citizens would create a capital of 10 million centimes; that is a nice lever for moving mountains."¹⁹³

The act of joining together thereby confers a certain value in itself, enabling what is not possible for individuals in isolation. On the theory of moral expectation, such added value could justify paying more for insurance than one's actuarial premium. It also suggests that the individual has an interest in the survival or wellbeing of the association beyond whatever strict financial benefit he expects to derive. Quetelet influentially pointed this out, noting that those who insure themselves "present to the state a type of guarantee that they respect the public order. They will not, in effect, compromise the future of their families, in exposing without reason the product of their labors to the chances of political dislocations." Someone who is literally willing to pay a

¹⁹² See also François Ewald, *L'État providence* (Paris: Grasset, 1986), 183-5.

¹⁹³ Edmond About, *L'assurance* (Paris: 1865), 35, quoted in Ewald, *L'État providence*, 183.

premium for his security will not let potentially disruptive opinions or passions get in the way of his peace and tranquility. Indeed, Quetelet continued, “I am often surprised that governments do not take a more direct part in institutions that can develop so advantageously the spirit of order and the morality of a nation.”¹⁹⁴ Napoleon III and, later, Bismarck would take heed in devising their own social insurance programs.

The idea of a moral duty to insure thus goes hand in hand with an attachment to or concern for the collective as such. Francois Ewald identifies in this logic one source of the doctrine of social solidarity, positing an obligation of each on behalf of all.¹⁹⁵ This view has much to recommend it, particularly given later 19th- and early 20th-century thought defending social insurance in solidaristic terms.¹⁹⁶ Yet insofar as the late-classical account of insurance still rested on an expectation calculated for single events, and therefore on the exchange of a privately owned risk, it retained a fundamentally individualist methodology and orientation. Insurance on this view remains a contract for the quantifiable benefit of the individual insured, even if that benefit extends beyond his strict mathematical expectation.¹⁹⁷ The state can encourage citizens to act in their interest or facilitate the pooling of their risks, but the social order and whatever solidarity it inspires are ultimately a means to personal, possessive ends.

¹⁹⁴ Quetelet, *Lettres*, 47.

¹⁹⁵ Ewald, *L'État providence*, 185.

¹⁹⁶ For several examples from the political realm, see T. H. Marshall, “The Welfare State—A Comparative Study,” in *Sociology at the Crossroads and Other Essays* (London: Heinemann Educational Books, 1963), 303-4.

¹⁹⁷ To take just one example, in 1837 Poisson defined mathematical expectation as the *individual's* due portion of a sum that can be imagined as shared between parties to an aleatory contract. He continues, “If the gain is 60,000 francs, for example, and 1/3 is the chance of the event to which it is attached, the person who is to receive that sum eventually could consider a third of 60,000 francs as a good that he possesses, and which ought to be included in the list of his actual fortune.” Poisson, *Recherches*, 71-2.

B. Causal laws and rational planners

If the liberal side of the liberal-statist account of social insurance was its assumption of an individually calculable expectation, supplemented for the purposes of decision-making by a utility value, its statist side emphasized statistical observation and aggregate regularities. From the elevated vantage point of the social planner, Quetelet explained, “moral phenomena...resemble physical phenomena,” generated by the causal forces or “general facts, by virtue of which society exists and is preserved.”¹⁹⁸ Each individual, “as a member of the social body...is subjected every instant to the necessity of these causes, and pays them a regular tribute...”¹⁹⁹ By examining aggregate data from savings banks and insurance companies, for example, the social physicist can “ascertain the degree of foresight at different periods of life,” and understand what conduces to the development of those “virtues most essential to the social state.”²⁰⁰ Enlightened leadership can then alter social conditions in such a way as to, on average, encourage such virtues and minimize their corresponding vices.²⁰¹

Once individual misfortunes and even vices can be explained as a product of social forces, which can be altered at will by a “few men, gifted with superior genius,” it would seem that the latter have an obligation to modify or at least mollify conditions on behalf of those who suffer from them.²⁰² Quetelet continued,

“...[S]ince the crimes which are annually committed seem to be a necessary result of our social organization, and since the number of them cannot diminish without the

¹⁹⁸ Adolphe Quetelet, *A Treatise on Man and the Development of his Faculties* (Edinburgh: William and Robert Chambers, 1842), 6.

¹⁹⁹ *Ibid.*, 7.

²⁰⁰ *Ibid.*, 78.

²⁰¹ In this respect Quetelet parted ways with Condorcet, who as Keith Baker points out prioritized individual choice over social engineering. Yet Condorcet was very much a progenitor of Quetelet’s statistical approach, one that effectively reduced the social sphere to the state and a mass of individuals, understood as “quantifiable and interchangeable units.” Baker, *Condorcet*, 262.

²⁰² Quetelet, *Treatise*, 6.

causes which induce them undergoing previous modification, it is the province of legislators to ascertain these causes, and to remove them as far as possible... Indeed, experience proves as clearly as possible the truth of this opinion, which at first may appear paradoxical, viz., that *society prepares crime, and the guilty are only the instruments by which it is executed*. Hence it happens that the unfortunate person who loses his head on the scaffold, or who ends his life in prison, is in some manner an expiatory victim for the society. His crime is the result of the circumstances in which he is found placed...²⁰³

This is a version of the same argument that motivated industrial insurance and workers' compensation laws, the first frontier of the modern welfare state, in France and other countries through the second half of the 19th century. Once accidents are seen as an objective fact or inevitability, outside the sphere of individual agency, it is possible to treat those whom they befall as innocent victims of economic progress or the common good. Indeed, on this view one can derive a collective responsibility to the individual whose unlucky sacrifice enables the entire system to continue.²⁰⁴

If insurance is a social duty, however, does this mean it should be compelled? This was a controversial question in the years leading up to and after the Revolution of 1848, in part because insurance had emerged within a basically liberal paradigm, and despite its increasingly aggregative character it had not yet shed the image of a voluntary exchange. It could also be said that the argument from personal obligation, insofar as it aimed to cultivate bourgeois virtues such as saving, foresight, and familial responsibility, aspired to reflect a choice on the part of workers themselves. While the case for social duty on one hand denied individual agency in causing or preventing accidents, then, on the other it still rested on a view of prudence as foreseeing calamity and taking steps to mitigate its effects.

²⁰³ Ibid., 108 (italics original).

²⁰⁴ Insurance as a social obligation in this sense looks not unlike compensation-based policies designed to recognize military service and other contributions to public life that began to emerge at roughly the same time. For the American case, see Theda Skocpol, *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States* (Cambridge and London: Harvard University Press, 1992).

In France, this debate played out in discussions of the country's first social insurance laws, adopted by the legislative assembly in the summer of 1850.²⁰⁵ The first of these was a national pension fund, designed to channel workers' retirement savings, and the second concerned France's growing network of voluntary mutual-aid societies.²⁰⁶ The political controversy concerned whether pension payments should be obligatory. On one side stood the friendly society model, with its spontaneously communal spirit but patent limitations in serving the poor and the lower strata of the working class. On the other side stood the possibility of mandatory payments accompanied by sanctions for noncompliance. The republican legislature rejected compulsion in favor of greater freedom for the individual worker. But Napoleon III and his advisors sought to reinforce the role of the state, and therefore opted for a system of subsidized liberty, or helping those who help themselves. Individual savings and the interest rate they accrued would be guaranteed by the government, and the mutual societies would be recast as intermediaries between workers and the state.²⁰⁷ The aim was to generate on the part of individual laborers a sense of shared interest with society as a whole and an attachment to its representative, the providing (and provident) state.²⁰⁸ As one parliamentarian explained at the time, adopting the liberal position but echoing Quetelet, national insurance

²⁰⁵ Ewald avers that these laws may be seen, "without too much anachronism," as the first laws of French social insurance. Ewald, *L'État providence*, 208.

²⁰⁶ According to Émile de Girardin, the first two friendly societies in France were established in 1580 and 1694, and the third in 1760. The French revolutionaries set out to destroy them, along with corporations, and as a result only thirteen were formed between 1794 and 1806. Their number began to grow again after that, and by 1822 there were 132 in Paris alone. Girardin, *La politique universelle*, 371.

²⁰⁷ Anne Reimat, "Old Age Pensions, Demography, and Economic Growth in the Long Run: The French Case Revisited," *Historical and Social Research/Historische Sozialforschung* 37, no. 4 (2012): 329; Yves Saint-Jours, "France," in *The Evolution of Social Insurance 1881-1981*, ed. Peter A. Köhler and Hans F. Zacher (London: Frances Pinter, 1982), 112; and Ewald, *L'État providence*, 210.

²⁰⁸ Ewald, *L'État providence*, 208-213.

creates between the state and the worker a solidarity, a community of interests that cannot but profit the public peace.... The worker, once entered into this vast association, is interested for his entire life in the affirmation of society, in the development of public prosperity. His fortune is tied to that of the state.²⁰⁹

This position may have been the liberal one, but it also reflects the paternalism of Quetelet himself, putting faith in the power of enlightened planners to alter social conditions and with them (average) individual behavior. Indeed, Napoleon apparently saw in insurance a valuable means to his own overarching political goal of building a strong executive that could end the era of revolution by reducing economic discord and depoliticizing government.²¹⁰ If it is true that, as Ewald has argued, these mid-19th-century political developments “found their conditions of possibility in the formation of that probabilistic rationality” developed by Quetelet, it should come as little surprise that they also reflect the late-classical probabilists’ uneasy union of liberty and social control.²¹¹ Quetelet’s notorious “average man,” the target of proposed efforts at moral improvement, would over time come to represent an ever greater proportion of the population, as social divisions narrowed and everyone came to resemble that calm and orderly being, “alike removed from excess or defect of every kind.”²¹² Under such conditions, social insurance could be both equitable and encompassing, individually choiceworthy and collectively beneficial: The state would be one large mutual assurance association comprising thousands or millions of “average men.” In the meantime, some encouragement or compulsion might be necessary to show men their true interest and induce them to act on its behalf.

²⁰⁹ Ferrouillat, *Rapport au nom du Comité du travail*, Assemblée nationale, February 19, 1849, impression no. 896, 56, quoted in Ewald, *L'État providence*, 212-13.

²¹⁰ On Napoleon III's political vision and goals, see Roger Price, *Napoleon III and the Second Empire* (London and New York: Routledge, 1997), 9-10, 25.

²¹¹ Ewald, *L'État providence*, 219.

²¹² Quetelet, *Treatise*, 100.

C. The social insurance moment

These two early welfare laws, while clearly reflecting the distinct ambitions of Napoleon III himself, emerged at a time when the idea of social insurance had gained favor in diverse political circles in France. This is confirmed in three tracts from the years leading up to and during the Second Republic: G. S. Boyer's *Projet d'assurance générale de bienfaisance nationale et de secours mutuel dans les 86 départements*; Raoul Boudon's *Organisation unitaire and nationale d'assurance*; and Emil de Girardin's *La politique universelle*, all written between 1838 and 1852.²¹³ Their authors came from substantially different schools of thought: Boudon was a follower of utopian socialist Charles Fourier, while Girardin is best remembered as owner of the popular newspaper *La Presse*, which switched its political allegiance from conservative to republican between 1848 and 1852.²¹⁴ Viewed in conjunction, however, they point to the prominence of insurance rhetoric in mid-19th century French political discourse, precisely at the time when the foundations of that country's welfare state were laid. They also reflect the diverse distributive arrangements to which the late-classical approach, and its somewhat ambiguous duty to insure, lent itself in practice.

Boyer's scheme warrants only a brief note.²¹⁵ First published in 1838, it envisioned a single fund for each of France's 86 departments. The fund would comprise contributions from all types of worker, to be distributed to any member who incurred a qualifying event and upon retirement. Although Boyer proposed different contributions for men and women, and allowed for the

²¹³ G. S. Boyer, *Projet d'assurance générale de bienfaisance nationale et de secours mutuel dans les 86 départements*, 2nd ed. (Paris: 1842); Raoul Boudon, *Organisation unitaire and nationale d'assurance* (Paris : 1848); Girardin, *La politique universelle*. These three tracts are also the focus of the discussion in Ewald, *L'État providence*, 214-19.

²¹⁴ See Ewald, *L'État providence*, 214 and Christopher Guyver, *The Second French Republic, 1848-1852: A Political Reinterpretation* (London: Palgrave Macmillan, 2016), 5.

²¹⁵ Neither this work nor its author appears to have left much of a mark on posterity. The only secondary reference I have found to them is in Ewald, *L'État providence*.

formation of separate classes for urban and country-dwellers, he otherwise envisioned that participants would pay the same amount, with those who entered the fund later simply receiving proportionally less in benefits. The aim of this rough actuarialism, he explains, is to create an inclusive scheme in which each accident is everyone's concern, and "the entire popular mass becomes one's support in reciprocity."²¹⁶

Similarly, Boudon's proposal, published a decade later, explicitly rejects the idea of strict probabilistic justice. Boudon argues that too much risk classification in a national scheme would be administratively infeasible, and would only increase contributions for the most vulnerable.²¹⁷ Instead, he breaks the entire citizenry into five general risk categories, with payments for each group based on a single national average. He does, in the interest of fairness, suggest a practice of compensating those who face lower odds of the same accident with additional coverage to justify their equal payment. Yet his approach is deliberately approximate, eschewing precise probability calculations in favor of putting everyone into the same boat. "I insist greatly on this point because one should not be lead astray, like those who have come before, by the desire to arrive at a minutely exact distribution, to an abstract justice, that will entail considerable expenses without leading to that end..."²¹⁸

Such macro-level actuarialism, which applies probabilities to the system's overall finances but not to individual premiums, has two advantages on Boudon's account. First, it allows for tapping into large-scale averages and making the scheme secure for the individual in a way that more limited risk pooling might not. "The chances of probabilities are the more uncertain the more they are established in a more restrained circle. To the extent that one enlarges the base of operation,

²¹⁶ Boyer, *Projet d'assurance générale*, 7.

²¹⁷ Boudon, *Organisation unitaire*, 14, 19, and 43.

²¹⁸ *Ibid.*, 47.

probabilities become more and more regular, more and more favorable.”²¹⁹ In this respect social insurance is an eminently responsible way to pursue one’s own advantage: “The principle of all insurance consists in abandoning a minimum part of one’s revenue in the aim of assuring the conservation of capital.” As such it “preserves the members from the alteration or the loss of their fortune by means of a light sacrifice.”²²⁰

Second, macro-level actuarialism makes it possible to see insurance as a social obligation, a means by which the citizen body as a whole protects the faultless sufferers among them. Insurance is “eminently social; it interests the totality of insured in the conservation of the property of each one.”²²¹ Disaggregation by risk category will only create arbitrary distinctions and undermine the ultimate goal, “that all the population of a town will see itself as struck at the same time, lest an accident will be seen as a particular fortune.”²²²

Boudon is also clear, however, that not everyone is equally worthy of such a solidaristic embrace. In recommending an insurance fund for those “who, by productive labor, could be exposed to certain accidents and maladies,” he makes sure to contrast these deserving claimants from “the idle, who expose themselves voluntarily...to certain dangers without a single social utility,” and to whom the collective in return “owes nothing for individual accidents.”²²³ Whether intentionally or not, Boudon thus offers a useful reminder of the paternalistic and exclusionary side of social insurance. Even the most disaggregated risk pool will inevitably rest on a decision about who is, and who is not, considered a probabilistic equal for the purposes of mutual protection.

²¹⁹ Ibid., 24.

²²⁰ Ibid., 9.

²²¹ Ibid.

²²² Ibid., 36.

²²³ Ibid., 61-62.

The final and apparently most successful contribution to this movement of sorts was Girardin's *La politique universelle*. Girardin was a journalist and editor, who after launching a number of literary and political magazines in the 1820s and 30s went on to found the *Presse*, France's first popular newspaper.²²⁴ His biographer depicts him as a pioneer of modern journalism and originator of Napoleon III's ideology of liberal empire, a "day-before socialist" who imagined the future of government as technocratic, apolitical, and welfarist.²²⁵ He was at one point closely aligned with French economic liberals, though he also critiqued orthodox liberal positions—along with Napoleon III's early social insurance laws—as inadequate to the needs of the working population.²²⁶

In *La politique universelle*, Girardin envisioned "a society that, reducing everything mathematically to risks judiciously foreseen and to probabilities exactly calculated, has for its unique pivot universal insurance."²²⁷ Mathematical probability "applied to the life of nations, to the case of war and revolution, is the foundation of all high politics.... To govern is to foresee; to foresee nothing is not to govern, but to run to one's death."²²⁸ Among the practical proposals Girardin derived from this idea was a "right of work" (as opposed to a right *to* work), which would guarantee a minimum salary to workers and thereby ensure that they could provide for their own subsistence through labor. Under such a system, "individual foresight combined with universal insurance" will replace charity and governmental relief, and "Fraternity" will be transformed from a "rare and

²²⁴ Maurice Reclus, *Émile de Girardin: le créateur de la presse moderne* (Paris: Librairie Hachette, 1934).

²²⁵ *Ibid.*, 234, 200.

²²⁶ Émile de Girardin, *Le socialisme et l'impôt* (Paris: 1849) and *L'abolition de la misère par l'élévation des salaires: Lettres à M. Thiers* (Paris: 1850), 34-35, 40. See also Nathalie Sigot, "Utility and Justice: French Liberal Economists in the 19th Century," *European Journal of the History of Economic Thought* 17, no. 4 (2010): 759-792.

²²⁷ Girardin, *La politique universelle*, 23.

²²⁸ *Ibid.*, 19.

exceptional sentiment” into a “common and ordinary science.”²²⁹

According to Francois Ewald, the success of Girardin’s work came from its presentation of insurance as a way out of liberal individualism that did not fall into socialism, which harmonized liberty and mobility with equality and security.²³⁰ Indeed, Girardin appears above all to have expressed in popular form the overarching hope of late-classical probability—that the “exact science” of probabilities, penetrating to the causes of suffering, would at last replace division and revolution with a nonpartisan government of enlightened risk-management. In such a world, mystical evil becomes calculable risk, culpable wrongdoing becomes accidental harm, and a providential God becomes “L’Etat providence”—the foreseeing, calculating, and providing state.

Although these pamphlets were the work of practical, political men rather than philosophers of probability, they reflect the presuppositions and challenges of the theoretical approach considered throughout this chapter. Late-classical probability aspired to a rational social order in which individual choices would align with the common good. *A posteriori* likelihoods calculated on the basis of observation would be predictive in both the singular case and in the aggregate; they would guide individual reason with the same force, and to the same ends, as they would the initiatives of enlightened social planners. Mutual insurance embodied this promise of harmony between personal and collective benefit, the promise of a regime that turns self-interest into universal security.

In practice, however, as these works attest, such harmony rests on a rather limited account of both freedom and the common good. Even where the targeted risk group is broad, as in both Boudon’s and Girardin’s proposals, social insurance restricts its protections to the responsible and

²²⁹ Girardin, *L’abolition de la misère*, 50-51.

²³⁰ Ewald, *L’État providence*, 219.

deserving, defined as those who have taken steps to reduce their own risks. As Girardin put it, “As soon as I insure myself, I no longer need for someone else to relieve me; just as it would be wrong to relieve me if I made the mistake of not insuring myself.”²³¹ Insurance as a surrogate for virtue means that one cannot transgress the bounds of responsibility—defined by the counsels of *a posteriori* probability—and still maintain a claim to aid. This is fraternity “made a science,” its obligations precisely delineated and deviations curtailed. Protection is thereby limited to those who have relatively little need for it, those who have earned it by choosing to insure in the first place.

In effect, this is the vision I have attributed to Quetelet: an entire society of average men, equivalent and interchangeable in their propensities and vulnerability to misfortune. Too much deviation, too much freedom, means that individual choices might not align with the collective good, since those who see themselves as superior will prefer not to insure, while those who can’t muster the right amount of self-control will unduly burden the collective and lose their claim to relief. Like the principle of indifference, the most distinctive tool of late-classical probability, this approach seems to rest on a kind of epistemic and contractual freedom, a view that the individual can make sound calculations and rational decisions if left to his own devices. Yet in practice it conceals what are often arbitrary or relative judgments—determinations about which events are to be considered equally likely, or who is to be considered a probabilistic equal—behind a veneer of mathematical rigor.²³² Freedom thus invites paternalism, as the common sense of everyday

²³¹ Girardin, *L’abolition de la misère*, 54.

²³² This was a central critique of the principle in mid-19th century probabilistic thinking, beginning most prominently with George Boole’s 1854 *The Laws of Thought*, which showed that application of the principle of indifference could lead to contradictory results depend on the events selected as equally probable. For a helpful overview see John V. Strong, “John Stuart Mill, John Herschel, and the ‘Probability of Causes,’” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1* (1978): 31-41. Note also Keynes’s assertion that the “enunciation of this principle, as it is ordinarily expressed, cloaks, but does not avoid,” the “element of direct judgment” on the part of the calculator. Keynes, *Treatise*, 58.

induction cedes its authority to expertise.²³³

In the political realm, one can see this development playing out in the fight between mutual societies and the state that began with Napoleon III's reforms and continued throughout the development of centralized welfare provision.²³⁴ Far more than simply a battle over membership and resources, this conflict concerned the proper locus of decision-making authority, whether it ought to be the citizen or the expert, the private individual or the state.

As we will now see, the interpretation of probability that supplanted the late-classical one over the course of the 19th century offered a different solution to the difficulty of aligning individual judgments with statistical frequencies. This account is less individualist in its methodology and more democratic in its political expression. It also, I hope to show, left a strong but largely unacknowledged impression on the theory and practice of welfare in the 20th century and beyond.

²³³ Condorcet, for example, was quite explicit that rational public decisions could only be made by the scientifically enlightened few. See Baker, *Condorcet*, 243-44.

²³⁴ Despite Napoleon's decision to preserve and even support French mutualism, the societies themselves were openly politicized in opposition to him, particularly following his liberalizing reforms and during the electoral campaigns of 1869-70. Roger D. Price, *Napoleon III and the Second Empire* (London and New York: Routledge, 1997), 47. Napoleon saw mutual societies as an opportunity for social engineering, not for the exercise of individual liberty: "The mutual aid societies, as I understand them, have the precious advantage of reuniting the different classes of society, of causing an end to the jealousies that can exist between them... One gives thus to different communities an aim of emulation, one reconciles the classes and one moralizes individuals." See Bernard Gibaud, *De la mutualité à la sécurité sociale: conflits and convergences* (Paris: Les Éditions ouvrières, 1986), 37-8.

Chapter Three: The Collectivization of Risk and the Early Welfare States

The last chapter considered the origins of the idea of social insurance, particularly as it emerged at the end of the 18th century and gained force through the first half of the 19th. This period saw a growing concern for the implications of new economic realities, in particular their baleful influence on the working poor, reliant as they were on wage labor and vulnerable to any number of disruptions as a result. Actuarial mutual insurance, both of the commercial and the reformed friendly variety, promised to reduce this vulnerability through the newly theorized mechanism of risk pooling or spreading. The corresponding image of insurance as a fair and prudent agreement, which moreover shares the burden of uncertainty across a large population, paved the way for the idea that it could also be used on a political scale to achieve redistributive ends.

The probability calculus would on this account be the foundation of a rational politics, in which the counsels of educated individual judgment align with those of scientifically minded elites. The social mathematicians' optimism in this regard stemmed at least in part from a belief in the harmony of probability theory's epistemic and aleatory interpretations, as reflected in the principle of indifference and effectuated through moral expectation and the duty to insure. Thanks to this constellation of tools, late-classical probabilists were able to present insurance as at once individually choice-worthy and collectively just, liberal and egalitarian in nature.

The period that will be considered in this chapter—from the second third of the 19th century through the early 20th—saw an intensified political interest in social insurance to promote security for the working classes, cultivate a kind of social cohesion, and preserve capitalism despite the clear hardships of the poor. During this period, however, the philosophical argument for insurance changed, as the presumed harmony between the aleatory and epistemic interpretations of probability unraveled and the former prevailed, at least for a time. The decisive moment in this shift came in

the 1840s, with the advent of a new interpretation of probability as the empirical frequencies of observed events. Frequentism affirmed the value of insurance by insisting that calculation and prediction are possible solely with reference to a series or group. Yet it also changed the normative justification for the practice, replacing the idea of personal expectation with the class-based sharing of a wholly collectivized risk. Risk as a quantified and manageable likelihood could on this view no longer be the property of an individual, and the insurance contract therefore had to rest on the insured's identification with his reference class as a whole. That this class had a real but relative existence, capable of being defined and redefined according to the available evidence, put insurance on a foundation of what might be called pragmatic solidarity, indispensable yet fluid in nature. As it happens, this vision accords very well with the way in which the first social insurance programs developed, as the result of lobbying by or on behalf of changeable aleatory classes. If the early welfare state marked a departure from liberal ontology and its accompanying morality, then, contemporary developments in probability theory can help to illuminate some of assumptions and implications of that shift.

I. The rise of the collective view of chance

A. Frequentism and its modified case for insurance

The writings of early statistical enthusiasts such as Condorcet and Laplace had treated mutual insurance as a predominantly voluntary affair, albeit one that would in some cases be administered on a political scale. Underlying this image of insurance as a voluntary contract was the assumption of an individual entitlement and the personal probability value on which it is based. Yet while late-classical probabilists continued to affirm the justice of insurance on grounds of individual entitlements, they simultaneously paved the way for a different approach, one in which probabilities

apply to groups rather than individuals and the benefits of insurance accrue primarily on a collective level. As Quetelet put it, “although the tables of mortality teach us no direct application to an individual, yet they offer very certain results when applied to a great number of persons.”²³⁵ Insurance came to exist somewhere between an individual choice and a social responsibility, a private affair and an investment in the state as the ultimate insurer.

With the emergence of the frequency theory of probability in the 1840s, the late-classical account of probabilistic justice suddenly faced a serious philosophical challenge. The new approach, picking up on the statistical orientation of late-classical probability, argued that all probabilities are simply long-run frequencies, derived from empirical observation. Frequentism therefore emphasized the difference between probability’s aleatory and epistemic interpretations, putting its weight decisively on the former, and in so doing also called into question the original contractual apparatus of classical probability. On a frequentist account, a probability value belongs exclusively to a series and has no meaning or existence in the context of individual events. Frequentism therefore entailed not only epistemological holism, according to which knowledge of the individual case derives from knowledge of the collective, but also a kind of ontological holism, implying the priority of the series as a matter of empirical fact. As the prominent frequentist John Venn put it, “Such regularity as we trace in nature is owing, much more than is often suspected, to the arrangement of things in natural kinds, each of them containing a large number of individuals.”²³⁶ In assigning a mathematical expectation to any individual, then, the frequentist intends “nothing more than to make a statement about the average of his class.”²³⁷ This emphasis on the priority of the

²³⁵ Quetelet, *Treatise*, 7.

²³⁶ John Venn, *The Logic of Chance*, 3rd ed. (New York: Dover, 2006), 55.

²³⁷ *Ibid.*, 151.

series or statistical class had important implications for the justification of insurance, which by then was widely seen as one of the most important practical applications of mathematical probability.

Frequentism was initially worked out during the late 1830s and early 1840s, during a rising tide of philosophical empiricism in England, France, and elsewhere. In 1843, four authors in three countries—Robert Leslie Ellis and John Stuart Mill in England, Antoine Augustin Cournot in France, and Jakob Friedrich Fries in Germany—proposed coherent versions of the theory, all arguing that probabilities ought to be interpreted not as measures of partial belief but rather as actual likelihoods derived from physical observation. As Ellis put it, “When on a single trial we expect one event rather than another, we necessarily believe on a series of similar trials the former event will occur more frequently than the latter.”²³⁸ Our judgments of probabilities are therefore derived from observed frequencies and depend not on the “fortuitous and varying circumstances of each trial” but rather on the natural fact that “on the long run, the action of fortuitous causes disappears.”²³⁹ Or, in the words of Venn, who sought to avoid causal assumptions completely, probability is concerned with those classes of things “as to the individuals of which we feel quite in uncertainty, whilst as we embrace larger numbers in our assertions we attach greater weight to our inferences.”²⁴⁰ Although the old view of probability as partial belief remained in force in some circles, these authors all took the regularities of social statistics and insurance as modeling an alternative view, in which the subjective or epistemological side of probability is “a mere appendage of the objective” or aleatory, affording “in itself no safe ground for a science of inference.”²⁴¹

²³⁸ R. L. Ellis, “On the Foundations of the Theory of Probabilities,” *Transactions of the Cambridge Philosophical Society* 8, no. 1 (1849): 1 [read February 14, 1842].

²³⁹ *Ibid.*, 3.

²⁴⁰ Venn, *Logic of Chance*, 3. While Cournot had maintained that probability should seek to “penetrate as much as possible into the knowledge of the thing in itself,” Venn and others insisted that statistics precisely evaded causal knowledge. A. A. Cournot, *Exposition de la théorie des chances et des probabilités* (Paris: L. Hachette, 1843), 185.

²⁴¹ Venn, *Logic of Chance*, 138.

Frequentism was not a unified movement, and its exponents expressed a range of views regarding the aims of the probability calculus and its relationship to psychology or inductive reason.²⁴² Nevertheless, what broadly united their program was a deliberate break from Laplacean probability, and in particular a critique of the principle of insufficient reason. As we saw, Laplace and his immediate successors had invoked this principle—which states that under conditions of ignorance, it is reasonable to assign equal probabilities to all of the possible outcomes of an event—to quantify likelihoods for which there is no clear empirical evidence. To this idea, frequentists responded that an assignment of equal probabilities must express an expectation about long-run outcomes, and that where one has no such information, “it is not easy to see how we are to justify any other opinion or statement...than a confession of such ignorance.”²⁴³ This is particularly true of individual events, since empirical probabilities are calculated for a sequence of trials each of which is entirely unpredictable. For any single event there cannot be a partial belief—or at least not one that corresponds to any actual state of the world—because probability values furnish no guide whatsoever to the likely outcome.

Many frequentists, most famously Mill, mollified over time their critique of the principle of insufficient reason, particularly as applied to causal hypotheses.²⁴⁴ Yet by insisting on the inapplicability of empirical probabilities to individual events, they had already planted the seeds of a radical critique of the classical rationale for insurance. Venn, for example, explained that from an

²⁴² Ellis, for example, saw his project as aligning probability “with a philosophy which recognizes ideal elements of knowledge,” in particular a belief “in the general similarity of the past to the future,” and “which makes the process of induction depend on them.” Ellis, “On the Foundations of the Theory of Probabilities,” 3, 6. Venn, meanwhile, set out to undermine completely “the conception of the science of Probability as a science of the laws of belief.” See John Venn, *The Logic of Chance*, 4th ed. (New York: Chelsea, 1962), 138, quoted in Philip Mirowski, “Marshalling Unruly Atoms: Understanding Edgeworth’s Career,” in *Edgeworth on Chance, Economic Hazard, and Statistics*, ed. Philip Mirowski (Lanham, Maryland: Rowman & Littlefield, 1994), 46. (References hereinafter to Venn’s *Logic of Chance* will be to the third edition.)

²⁴³ Venn, *Logic of Chance*, 142.

²⁴⁴ See Mill, *A System of Logic, Ratiocinative and Inductive*, Ch. xviii, in *Collected Works of John Stuart Mill*, vol. 7, ed. J. M. Robson (Indianapolis: Liberty Fund, 2006): 534-547 and compare the first (1843) edition, in *ibid.*, vol. 8, 1140-1150.

individual perspective insurance cannot be justified for any single event.²⁴⁵ Cournot allowed for the use of probabilities “to fix the conditions of a bet or some random business deal,” but made clear that whatever number the parties happen to choose is simply a “subjective” assessment, which will differ depending on each one’s knowledge and has nothing to do with any actual ratio of outcomes.²⁴⁶ Even Mill continued to maintain in later editions of his *Logic* that empirical averages, while necessary for practical use, “are of extremely small value as grounds of expectation in any one individual instance.”²⁴⁷

If mathematical probability is of such little use as a guide to single events, what if anything is left of the justification for insurance? Venn was the first to clearly recognize that for insurance to remain defensible on an individual level, it would require a different account from the one offered by classical probability. One possible approach, he suggested, is for the individual to consider his own actions as a series, and to find that the “equalization of his gains and losses, for which he cannot hope in annuities, insurances, and lotteries taken separately, may yet be secured to him out of these events taken collectively.”²⁴⁸

This approach is problematic, however, because it’s unclear what actions constitute the relevant trials and how they can be considered a series, which Venn defines as a set of events that are completely uniform with regard to most of their qualities.²⁴⁹ A more promising line is Venn’s

²⁴⁵ Venn, *Logic of Chance*, 151, 148.

²⁴⁶ Cournot, *Exposition de la théorie*, 438; cf. 160, 183-4. Similarly Boole, although not a frequentist, insisted that the “rules which we employ in life-assurance, and in the other statistical applications of the theory of probabilities, are altogether independent of the *mental* phaenomena of expectation. They are founded upon the assumption that the future will bear a resemblance to the past; that under the same circumstances the same event will tend to recur with a definite numerical frequency; not upon any attempt to submit to calculation the strength of human hopes and fears.” George Boole, *An Investigation of the Laws of Thought* (New York: Dover, 1854), 244.

²⁴⁷ Mill, *System of Logic*, 542, 538.

²⁴⁸ Venn, *Logic of Chance*, 148. Condorcet had intimated a similar solution in his own essay on maritime insurance, but unlike Venn he continued to regard probability values as relevant to individual trials.

²⁴⁹ *Ibid.*, 7.

suggestion to “suppose the existence of an enlarged fellow-feeling,” such that each man perceives “that the gain of others offers some analogy to his own gains.”²⁵⁰ The justice of insurance would thus hinge on each person’s ability to see himself first and foremost as a member of his class and to enlarge his own interest to encompass the group as a whole.²⁵¹ This class-based actuarialism is grounded in a kind of solidaristic sympathy rather than an individual entitlement. It is also flexible, based on an admission that the insured’s designated reference class—those with whom he is considered to be equal for actuarial purposes—can vary according to the insurer’s needs and available information.²⁵² On this view, the class exists in nature as a class, but because it is known through observation and not as a consequence of any deeper causal account, there is no reason to fix its existence as such or to consider it as inherently superior to any other observed series.

Frequentism also suggested a new interpretation of probabilistic equality, based on its understanding of the nature of chance itself. Since Hume, at least, the idea of chance was associated with a mental state of complete indifference between the possible outcomes of a particular trial.²⁵³ Frequentists, by contrast, in keeping with their thoroughgoing empiricism, looked for chance in the world and found it in the observation of causal independence or the absence of a discernable agency.²⁵⁴ We can know that an outcome is chancy on this view not because we are ignorant about

²⁵⁰ Ibid., 149.

²⁵¹ Interestingly, the word “solidarity” appeared in the English language in 1841, shortly before the emergence of frequentist probability theory, having come directly from the French. See “Solidarity,” *Oxford English Dictionary*, <http://www.oed.com.ezp-prod1.hul.harvard.edu/view/Entry/184237?redirectedFrom=solidarity#eid>. In 1843 Cournot used the word in a technical sense, to refer to causes that are not independent of one another and therefore produce outcomes that cannot be considered random. “To properly understand [the notion of randomness], one must attach to it only what is fundamental and categorical in the notion of chance, the idea of the independence or the absence of solidarity between different series of events or causes.” Cournot, *Exposition de la théorie*, 77.

²⁵² Venn, *Logic of Chance*, 224-231.

²⁵³ See S. L. Zabell, *Symmetry and its Discontents: Essays on the History of Inductive Probability* (New York: Cambridge University Press, 2005), 18-26, for a brief but illuminating survey of views about equiprobability and chance from Abuthnot to Frank Ramsey.

²⁵⁴ On randomness and causation, see Venn, *Logic of Chance*, ch. 5 and Cournot, *Exposition de la théorie*, ch. 4.

its result beforehand but because repeated trials of the same event generate a uniform distribution in which the individual results are random, or occur with equal frequency. Although the distinction is a subtle one, the frequentist view implies that we can learn from experience that any given risk—say, a work accident or losing one’s job—is equally distributed within a specific population in the same way that natural events such as deaths are.²⁵⁵ Many of the earliest welfare policies rested on precisely such claims about the revealed randomness of particular phenomena.

B. Frequentism and utilitarianism

One would not wish to overstate the influence of frequentist probability theory on ideas about social solidarity or on early welfare policy itself. Except perhaps in England, frequentism never gained wholesale acceptance within the philosophy of probability, and it was therefore somewhat limited in its theoretical influence.²⁵⁶ Venn’s rejection of the principle of indifference in particular, and of the application of numerical probabilities to individual events, met with skepticism even among contemporaries who were sympathetic to his empirical approach.²⁵⁷ Indeed, it is difficult to make out a compelling case that any given contribution to the rarefied debates within probability theory had a direct influence on the realm of social or political practice.²⁵⁸

²⁵⁵ Venn, Peirce, and Edgeworth all affirmed this account of randomness. For a summary and critique of their views, see Keynes, *Treatise*, 331-2.

²⁵⁶ According to Keynes, writing in 1921, frequentism had a “large measure of...acceptance in England at the present time,” thanks to the fact that most English writers had approached probability “from the statistical side.” *Ibid.*, 121.

²⁵⁷ On the partial rejection of Venn’s critiques by the giants of statistics at the time, Edgeworth and Pearson, and his limited influence outside of England, see Zabell, *Symmetry*, 121-23; cf. F. Y. Edgeworth, “The Philosophy of Chance,” *Mind*, New Series 31, no. 123 (1922), 261 (hereinafter “The Philosophy of Chance (II)”).

²⁵⁸ Moreover, there appear to be other, more direct sources of the philosophical holism that emerged at the end of the 19th century and arguably supported early welfare legislation by promoting ideas about solidarity and social obligation. See, e.g., Andrew Vincent and Raymond Plant, *Philosophy, Politics and Citizenship* (Oxford: Basil Blackwell, 1984), 47-48; Michael Freedman, *The New Liberalism: An Ideology of Social Reform* (Oxford: Clarendon Press, 1978), 18; and Peter Baldwin, *The Politics of Social Solidarity: Class Bases of the European Welfare States, 1875-1975* (Cambridge: Cambridge University Press, 1990), 35. On the role of idealism in limiting Venn’s immediate influence, see Zabell, *Symmetry*, 123.

Nevertheless, there are a number of indirect indications that a roughly frequentist understanding of probabilities helped to undergird early social legislation. The first is that the flexible class-based actuarialism of the frequentist account does in fact characterize the shape of many early welfare states, as I will describe in greater detail below. Second, the frequentist subordination of causal inquiries to the search for observed large-scale regularities supports the collective understanding of risk and responsibility that was so important in justifying social insurance, and unlike late-classical accounts it does so without simultaneously invoking an individual entitlement. Third, the demotion of causal analysis also had the effect of downplaying individual responsibility and thus may have supported the case for compulsory insurance, since on this view the benefits of risk pooling hinge on its collective effects and not on whatever personal virtues it may reflect. Along the same lines, the frequentist account of aleatory solidarity calls for an identification of the one with the many that is unlikely to arise spontaneously and may therefore have to be compelled for the true benefit of the insured, as early social insurance so often was.

These are all indications of what one might call an “elective affinity” between the philosophy of probability and early welfare thinking. Yet there is also a direct historical connection of some significance. Utilitarian philosophers and political economists of this time, particularly in England, took considerable interest in the foundations of probability, and there was a deep affinity between their theoretical commitments and the frequentist view. Both took an aggregative approach to individuals in the name of a common or collective good, and both rested on an abstract assumption of equality that allowed for such aggregation. In addition, utilitarian skepticism about natural rights finds a direct parallel in the frequentist rejection of a personalized risk, reflecting in each instance an explicit departure from classical liberal thought.²⁵⁹

²⁵⁹ According to some historians, responsibility for the intellectual foundations of the early British welfare state lies not with the Fabian society—Lord Beveridge himself was never a Fabian—or any other socialist school but rather with the economic theories of the Cambridge Marxallians and J.A. Hobson. A major question that occupied these thinkers,

Venn, for one, recognized the link between his interpretation of probability and utilitarianism. In his final addition of the *Logic*, he recommended the latter as the successor to and fulfillment of the concept of moral expectation, in that it answers the question of which “distribution of wealth tends to secure the maximum of happiness.”²⁶⁰ As we have seen, late-classical probabilists invoked moral expectation to explain why the dictates of mathematics do not always comport with those of common sense. In particular, they justified the immorality of gambling, in contrast to the virtue of insurance, on the ground that whatever loss the gambler may incur will produce in her more pain than the pleasure she would derive from the same amount should she win. Venn accepts this basic assumption about the concavity of utility curves. Yet whereas earlier authors had used moral expectation to justify individual decisions about risk, Venn interprets it as an argument about the overall distribution of wealth. If the disutility of the losing gambler exceeds the utility of the winner, then overall happiness has decreased, and what is proved is that “*inequality* is bad, on the ground that two fortunes of £50 are better than one of £60 and one of £40.”²⁶¹ The real problem with gambling, therefore, is not its disutility to the individual but “its tendency to the increase of the inequality in the distribution of wealth,” a conclusion that recommends the “Socialist’s ideal as being distinctly that which tends to increase happiness.”²⁶²

It was, in particular, the utilitarianism of his contemporary Frances Ysidro Edgeworth that Venn cited as the successor to moral expectation. The reference, although made to Edgeworth’s chagrin, is nevertheless revealing of the intersection between probability theory and political

who include Alfred Marshall, Edgeworth, and Arthur Pigou, was how to envision social consensus that is not simply an aggregation of individual evaluations. Here Hobson’s account of the organic nature of society has been cited as helping to solve the problem by allowing for a shared conception of social need—a project to which frequentism, I wish to suggest, lent strong if indirect support.

²⁶⁰ Venn, *Logic of Chance*, 392-93.

²⁶¹ *Ibid.*, 390.

²⁶² *Ibid.*, 391, 392.

economy in the latter half of the 19th century.²⁶³ If one were to try to trace a direct historical connection between the frequentist account of likelihoods and the interventionist utilitarianism of the early British welfare state, it is clear that Edgeworth—who used statistical analogies to reimagine market economics; who was a close associate of Sydney and Beatrice Webb; and who late in life became the intellectual target of young John Maynard Keynes, who shared many of his interests and influences—would play an important role. Although this is not the place to fully trace that intricate lineage, it is worth pausing to consider his work.

Edgeworth was not a conventional frequentist, but like Venn he was an empiricist when it came to probability, insisting that likelihoods should “rest upon precise experience” if they are to be measurable at all.²⁶⁴ While he defended the principle of indifference, and was even responsible for persuading Venn to moderate his critique of inverse probability in later editions of the *Logic*, he did so on grounds that were substantially different from those offered by many of his late-classical predecessors. “In Probabilities,” he explained, “it is often necessary to assume that quantities between which no inequality has obtruded itself in the course of experience may be treated as

²⁶³ An extended discussion of their relationship and points of dispute can be found in Mirowski, “Marshalling Unruly Atoms,” 46-47. Mirowski’s account of the development of Edgeworth’s statistical thought is the source of my suggestion that Edgeworth sought to reconceive economics along stochastic lines. Here Mirowski claims that Edgeworth was particularly provoked by Venn’s invocation of him as an authority against the practical relevance of moral expectation.

²⁶⁴ F. Y. Edgeworth, “The Philosophy of Chance,” *Mind* 9, no. 34 (1884): 235 (hereinafter “The Philosophy of Chance (I)”). It was Keynes who explicitly situated Edgeworth along with Ellis, Venn, and others in the frequentist camp, as part of a strategy to replace their views with his own account of probability as a logical relation. See Mirowski, “Marshalling Unruly Atoms,” 51. Yet Keynes also had good reason for this designation, and Edgeworth continued to indicate even later in life that the probability calculus is grounded in physical law, suggesting for example that “the principles of Probabilities” may “owe something to the principles of mechanics which rule the movement of molecules, which underlies the phenomena of chance.” F. Y. Edgeworth, “On the Use of the Theory of Probabilities in Statistics Relating to Society,” *Journal of the Royal Statistical Society* 76, no. 2 (1913): 174. Theodore Porter has characterized Edgeworth as maintaining a “subjectivist interpretation” of probabilities, “albeit a more nuanced one” than his contemporary Stanley Jevons, but I find little evidence for this conclusion, except perhaps if by “subjectivist” it is meant someone who is interested in the relationship between objective frequencies and individual belief. Theodore Porter, *The Rise of Statistical Thinking: 1820-1900* (Princeton: Princeton University Press, 1986), 97.

equal.”²⁶⁵ Despite his deference to authority, however, Edgeworth emphasized that this assumption rests “upon the loose foundations of common-sense,” not on the statistician’s mental state of uncertainty.²⁶⁶ Specifically, there is a “substratum established by wide experience, that what has held good in two or three...experiments will hold good generally.” Equiprobabilities based on such observation will therefore be “respectably grounded upon experience,” and not “so inane as Mr. Venn would have us believe.”²⁶⁷

Edgeworth explicitly related this commonsense assumption of equal probabilities to utilitarian ethics, explaining that in the latter “Equality is similarly postulated. The reasoning of Bentham and Prof. Sidgwick, that equality of distribution tends to maximum happiness, presupposes that the distributees are equally capable of happiness.”²⁶⁸ Indeed, the application of “intellectual probability...appears to be unconsciously performed by the utilitarian who thinks it ‘fair’ to treat as equals those between whom no material difference is discerned.”²⁶⁹ The assumption of equality thus serves the practical needs of the utilitarian calculus in the same way that it serves those of scientific endeavor: It provides “an hypothesis which may serve as a starting point for further observation” and calculation.²⁷⁰

²⁶⁵ F. Y. Edgeworth, “The Method of Measuring Probability and Utility,” *Mind* 12, no. 47 (1887): 484; cf. his “Applications of Probabilities to Economics,” *The Economic Journal* 20, no. 78 (1910): 284-304.

²⁶⁶ Edgeworth, “Probability and Utility,” 484.

²⁶⁷ Edgeworth, “The Philosophy of Chance (I),” 234-35.

²⁶⁸ Edgeworth, “Probability and Utility,” 484. Sidgwick states this explicitly in his *The Methods of Ethics* (London: Macmillan & Co., 1874), 387.

²⁶⁹ Edgeworth, “The Philosophy of Chance (I),” 233.

²⁷⁰ Ibid.

Eventually, such observation and calculation would allow for a truly “scientific hedonimetry,” a project that Edgeworth first laid out in his 1881 *Mathematical Psychics*.²⁷¹ There he analyzed the phenomenon of contractual indeterminacy—that is, the fact that in the absence of perfect competition, there is more than one desirable outcome for any exchange—as the fundamental problem of economic and social life. Under such conditions, either party to an agreement may choose to re-contract after the initial bargain is made, giving rise to a general state of uncertainty. The advent of trade unions and cooperative associations, while rational and beneficial to the participants themselves, exacerbates the problem, undermining competition and reducing the number of actors within a given field. Assuming then that individuals prefer certainty to uncertainty, “there would arise” from this situation “a general demand for a principle of arbitration,” a single solution for each type of contract that will ensure the parties’ benefit and the overall social good.²⁷²

The central question motivating this analysis and, less explicitly, Edgeworth’s subsequent work in mathematical statistics is how to reconcile the free choices of individuals with the collectively beneficial or utilitarian outcome. In *Mathematical Psychics*, Edgeworth’s answer involves intervention by statistically informed planners who, with the help of data about individual preferences (to arbitrate contracts on the private level) and the characteristics of classes (to arbitrate on the social level), are able to facilitate socially advantageous agreements. Supplementing competition with arbitration in this way means that the “*economical* leads up to the *utilitarian calculus*,” and the outcomes of self-interested bargains can be fixed so as to promote overall utility.²⁷³ Over time, Edgeworth suggests, the application of scientific utilitarianism to economic and political life

²⁷¹ Edgeworth, *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences* (London: C. Kegan Paul & Co. 1881), 80-81. See also Mirowski, “Marshalling Unruly Atoms,” 34-46.

²⁷² Edgeworth, *Mathematical Psychics*, 51.

²⁷³ *Ibid.*, 54-6.

will allow for a “reconciliation between *egoism* and *altruism*,” between a contractual regime guided by self-interest and one in which each citizen identifies his own good with that of the whole.²⁷⁴

It bears emphasizing that while this vision is thoroughly moral in its ambitions, there is nothing inherently egalitarian about it.²⁷⁵ Just as the principle of indifference, with its assumption of equal prior probabilities, would ultimately give way to empirical frequencies, the utilitarian assumption of equality could be modified in light of further observation, revealing that particular groups of people are more or less capable of happiness. A truly scientific utilitarianism calls for estimating “the duration of the pleasure, the susceptibility, as well as the number” of all citizens on the basis of extensive observation.²⁷⁶ It also calls for a rejection of those “metaphysical ‘rights of man’” often invoked to justify equality.²⁷⁷ The social contract on Edgeworth’s account is therefore to be informed not by natural rights but by an awareness of statistical groups and their potentially distinct distributive claims—a view that echoes what I will call the ethical-collectivist case for social insurance that emerged in conjunction with frequentist probability.

Although Edgeworth would eventually modify his youthful anti-egalitarian stance, he continued to emphasize the analogy between probability theory and utilitarianism, particularly as he turned to theoretical statistics in the decade after publishing *Mathematical Psychics*.²⁷⁸ For one, the

²⁷⁴ Ibid., 102.

²⁷⁵ As Mirowski also notes, quoting Edgeworth’s 1877 *New and Old Methods of Ethics*, its young author was “most anxious to reconcile his reverence for demonstrable innate superiority with his utilitarianism,” which in this instance he set out to do using the “language of thresholds and sensitivities” borrowed from German physicist and mystic Gustav Fechner. Mirowski, “Marshalling Unruly Atoms,” 15.

²⁷⁶ Edgeworth, *Mathematical Psychics*, 80-81; cf. 102: “It may be replied, greater uncertainty of hedonimetry in the case of others’ pleasures may be compensated by the greater number of measurements, a wider average; just as, according to the theory of probabilities, greater accuracy may be attained by more numerous observations with a less perfect instrument. The proposition, ‘the exercise of higher intellect is accompanied with greater capacity for pleasure,’ is proved by taking a wide average rather than by the self-observation, however accurate, of a single, perhaps exceptional, individual.”

²⁷⁷ Ibid., 129.

²⁷⁸ See, e.g., F. Y. Edgeworth, “*The Labour Party’s Aim: a Criticism and Restatement*, by seven members of the Labour party,” *The Economic Journal* 33, no. 132 (1923): 541.

analogy allowed him to affirm that utility values are measurable, since probability and in particular the normal law provided a means of approximating numerical measurements even where, as in the case of utilities, perfect measurement is impossible.²⁷⁹ In addition, and on a deeper level, the analogy allowed him to reconceive economic equilibrium as an evolutionary process, replacing both the simple self-interest of neoclassical theories and the imposed contracts of utilitarian arbitration with an account of stochastic order on the model of the gas laws.²⁸⁰ “The statistician finds in the world of atoms an ideal model of that law in which he should exercise himself day and night,” similar to the economist’s use of the theory of maxima as borrowed from mathematical physics.²⁸¹ The theoretical interest of the calculus thus lies in its ability to model the coexistence of free or spontaneous elements within a single harmonious system.²⁸²

This conclusion sheds light on why Edgeworth believed it was possible to replace both perfect competition and utilitarian arbitration as the regulating principle of economic life. In conjunction with this view, he also proposed an alternative psychological foundation for economic order. “Self-regarding self-interest, the gospel of Adam Smith, is not alone sufficient for industrial salvation: a leaf must be taken from his older and less familiar testament, of which the cardinal doctrine was *sympathy*.”²⁸³ Edgeworth is quick to clarify that by this he does not mean a “utopian eradication of self-love,” but rather “mutual understanding, between the parties to distribution.”²⁸⁴

²⁷⁹ See Edgeworth, “On the Use of the Theory of Probabilities,” 184.

²⁸⁰ Mirowski, “Marshalling Unruly Atoms,” 43.

²⁸¹ Edgeworth, “On the Use of the Theory of Probabilities,” 186.

²⁸² Ibid. and F. Y. Edgeworth, “Methods of Statistics,” *Journal of the Statistical Society of London* Jubilee Volume (1885): 194.

²⁸³ F. Y. Edgeworth, “The Theory of Distribution,” *The Quarterly Journal of Economics* 18, no. 2 (1904): 218, 216.

²⁸⁴ Ibid., 218.

Although he leaves the contours of this mechanism vague, one finds in such statements another strong echo of the frequentist case for social welfare, to which we will now turn.

C. Frequentist social welfare

The close connection between frequentism and utilitarianism for both Venn and Edgeworth—and, less explicitly, for Mill and Cournot—is not a coincidence. As we saw in the previous chapter, the very concept of utility, along with the assumption of the concave utility curve, emerged within the theory of probability. It was devised to reconcile mathematics with common sense, showing that probability is a reliable guide to decision-making as long as certain psychological and moral assumptions are incorporated into the calculus. Frequentism, too, set out to align probability calculations with “the common sense of mankind,” as Ellis put it, offering a more compelling basis for practical reason than its Laplacean predecessors.²⁸⁵ In doing so, it also pointed toward a more ambitious account of individual and collective welfare than classical liberalism had done.

As we saw in the last chapter, it was the assumption of ignorance at the core of late-classical probability that allowed it to remain liberal in its basic orientation, respecting (at least nominally) individual judgment about probabilities even when not based on empirical observation. Frequentism, by contrast, aspired to place its counsels on the firm ground of experience alone. It did this by rejecting any quantification of ignorance and by defining probability as relative to prior knowledge. From the individual’s perspective, this means that rational decisions cannot be made in isolation, but rather only through a deep identification with the experiences of others. Reason

²⁸⁵ Ellis, “On the Foundations of the Theory of Probabilities,” 5.

requires seeing one's own case as but one equally probable instance in an ongoing series of identical trials.

American pragmatist philosopher Charles Sanders Peirce was perhaps the most clear-eyed in recognizing this implication of the frequentist view of probability. Peirce explained that “in the long run, there is a real fact which corresponds to the idea of probability, and it is that a given mode of inference sometimes proves successful and sometimes not, and that in a ratio ultimately fixed.”

Because “probability essentially belongs to a kind of inference which is repeated indefinitely,” “there can be no sense in reasoning in an isolated case, at all.”²⁸⁶ Instead,

...logicality inexorably requires that our interests shall *not* be limited. They must not stop at our own fate, but must embrace the whole community. This community, again, must not be limited, but must extend to all races of beings with whom we can come into immediate or mediate intellectual relation.... Logic is rooted in the social principle.²⁸⁷

Peirce gives the example of choosing a card from a pack of 25 red cards and one black, or 25 black and one red. If choosing red will bring eternal felicity and black everlasting sorrow, one will certainly opt to pick from the first pack. Yet on Peirce's view there is no valid inference that justifies that choice. Because the exercise is only repeated once, there is no “real fact” whose existence gives truth to the statement that if he draws from one pack a particular color will likely appear. The example therefore illustrates that only by having enlarged interests—by caring “equally for what was to happen in all possible cases of the sort”—is it possible to act logically in choosing from the red pack.²⁸⁸

²⁸⁶ Charles Sanders Peirce, “On the Doctrine of Chances, with Later Reflections,” in *Philosophical Writings of Peirce*, ed. Justus Buchler (New York: Dover, 1955), 159-60.

²⁸⁷ *Ibid.*, 161-2.

²⁸⁸ *Ibid.*, 160-61.

The political consequences of this position are dramatic. Far from sanctifying the right of any individual to judge his own case, frequentism implies the epistemic priority of the community, and with it an argument for a kind of rational sympathy or altruism. Although this shift will not necessarily lead to different probability values, it does amount to a very different justification for both individual and state action. The truly reasonable person will ground his decisions in the “social principle,” caring for every other similar case in exactly the same way that he cares for his own.

Edgeworth had reached the same conclusion thirty years earlier. In analogizing the principle of indifference to the utilitarian assumption of equality, he was explicit that both must be founded on experience and revised in light of continued observation.²⁸⁹ The resulting empirically defined identity is what promotes the utilitarian ideal, wherein the individual “in a cool moment, [would] deliberately sacrifice his own greatest happiness to that of others.” This sacrifice is based on the decision maker’s understanding that the other’s wellbeing “counts for one” in comparison to his own, just as every event within a statistical series is identical to and carries the same weight as any other.²⁹⁰

Edgeworth’s vision thus bears a striking similarity to the “enlarged fellow-feeling” on which Venn’s account of insurance rests. Although “fellow-feeling” suggests an emotion, the ultimate basis of this disposition is rational rather than sentimental.²⁹¹ Specifically, it rests on the decision-maker’s knowledge about his own prospects, derived from repeated observation of those similarly situated, and on an understanding of the logical necessity of seeing himself as interchangeable with

²⁸⁹ Mirowski calls this “one of the more seductive analogies for Ysidro” in linking probability theory to utilitarianism. See Mirowski, “Marshalling Unruly Atoms,” 25-27, 40.

²⁹⁰ Edgeworth, *Mathematical Psychics*, 103.

²⁹¹ This is true for both Edgeworth and Venn. See, e.g., Edgeworth, “The Theory of Distribution,” 219, referring to “Intellectual sympathy.” Edgeworth’s utilitarian authority Sidgwick is also illuminating on this point: “... I find it very difficult to distinguish the sympathetic and the properly moral feelings in introspective analysis of my own consciousness: it seems clear that these two elements are continually combined, but it is hard to say precisely in what proportion.” Sidgwick, *Methods of Ethics*, 463.

everyone else in that group. As Edgeworth later put it, quoting himself in defense of the frequentist view, “‘A man, say, buys a life annuity, insures his life on a railway journey, puts into a lottery, and so on.’ It may be expected, I think, that the class of actions which cannot be regarded as part of a ‘series’ will diminish with the increase of providence and sympathy.”²⁹²

It would be a mistake to conclude from these examples that many 19th- and 20th-century thinkers understood the deeply collectivist implications of frequentism, let alone extrapolated from them a case for social insurance. Many of those now known as neoclassical economists, from Cournot to Alfred Marshall, held a roughly frequentist view of probabilities yet maintained rather conventional views about insurance.²⁹³ One partial exception is Arthur Pigou, whose remarks on the subject Edgeworth enthusiastically cited in a 1913 review.²⁹⁴ Pigou touted mutual insurance as a means for the “voluntary transference” of wealth from rich to poor, based on an empirically grounded sense of identity within particular groups of insured.²⁹⁵ Still, the epistemic and political ambition of the frequentist view—as well as its potentially non-liberal consequences—remained below the surface in many of these accounts. Instead, as I now hope to show, it was early welfare policy itself that brought it to the fore.

II. Risk in the early welfare states

There is no shortage of scholarly explanations for the rise of the modern welfare state, from industrialization and the destructive effects of capitalism, to efforts by non-democratic rulers to gain

²⁹² Edgeworth, “The Philosophy of Chance (II),” 260.

²⁹³ See, e.g., Alfred Marshall, *Principles of Economics*, 8th ed. (Hampshire and New York: Palgrave Macmillan, 2013), esp. 112, 332.

²⁹⁴ F. Y. Edgeworth, “*Wealth and Welfare*, by A. C. Pigou,” *The Economic Journal* 23, no. 89 (1913): 62-70.

²⁹⁵ Arthur Pigou, *Wealth and Welfare* (London: Macmillan, 1912), 366-69.

legitimacy, to the political power of working-class movements.²⁹⁶ Recent scholarship has shown a particular interest in the explanatory value of risk management and insurance as the motivations for social policies, with good reason.²⁹⁷ The idea of social insurance can be understood as a natural extension of early modern liberalism's preoccupation with security, particularly as guaranteed by contractual means, and of the statistical revolution that took place during and after the Napoleonic era, when, in Ian Hacking's words, "counting and measuring became the thing to do."²⁹⁸ The latter development encouraged recognition of the accidental and therefore undeserved character of many private misfortunes, and placed a greater onus on the collective to mitigate their effects. Historically and conceptually, then, ideas about risk have indeed been closely implicated in the development of many welfare regimes.

As fruitful as this recent scholarship has been, however, the literature has not seriously considered how changing understandings of probability and risk may have influenced the models of social insurance actually adopted. If, as I have been arguing, "risk" is not a unitary concept but one that has evolved, reflecting different normative commitments over time, then there is not a single insurance rationale for welfare but rather several, each embodying a different constellation of

²⁹⁶ For a fairly exhaustive list of explanations and an effort to empirically test each one, see David M. Cutler and Richard Johnson, "The Birth of the Social Insurance State: Explaining Old Age and Medical Insurance across Countries," *Public Choice* 120, no. ½ (2004): 87-121. The classic text in this field is Gösta Esping-Andersen, *The Three Worlds of Welfare Capitalism* (Princeton, NJ: Princeton University Press, 1990).

²⁹⁷ The recent literature, beginning with Francois Ewald's seminal *L'État providence*, in this vein is fairly large and rather diverse. Other noteworthy contributions include Baldwin's *The Politics of Social Solidarity*; Pierre Rosanvallon, *La nouvelle question sociale: Repenser l'État providence* (Paris: Éditions Du Seuil, 1995); Paul Johnson, "Risk, Redistribution and Social Welfare in Britain from the Poor Law to Beveridge," in *Charity, Self-Interest, and Welfare in the English Past*, ed. Martin Daunton (London: UCL Press Ltd., 1996), 225-248; Torben Iversen, *Capitalism, Democracy and Welfare* (Cambridge: Cambridge University Press, 2005); and Joseph Heath, "The Benefits of Cooperation," *Philosophy & Public Affairs* 34, no. 4 (2006): 313-351. For additional references to the insurance or risk-pooling rationale for welfare policies, see Philippe van Parijs, "Assurance, solidarité, équité: Les fondements éthiques de l'État providence," *Cahiers de l'École des sciences philosophiques et religieuses* 12 (1992): 49-71 and Michael Freedman, "The Coming of the Welfare State," in *The Cambridge History of Twentieth-Century Political Thought*, ed. Terrance Ball and Richard Bellamy (Cambridge and New York: Cambridge University Press, 2003), 5-44.

²⁹⁸ Ian Hacking, "Comment: In Praise of the Diversity of Probabilities," *Statistical Science* 5, no. 4 (1990): 452.

epistemic and moral claims. We have already seen that this is true in theory, through the differences between late-classical accounts of insurance and their frequentist successors. In what follows, I will attempt to show that it is true in practice as well. By examining some of the actual debates and policies surrounding the emergence of social insurance in Europe from roughly 1880 through the onset of the First World War, and in the United States up until the New Deal, one can see how reformers adopted, and adapted, the logic of probability to promote particular political aims.

I do not purport here to offer anything like an exhaustive explanation for these geographically and temporally diverse policies. I do suggest, however—and here my argument goes a step beyond previous work in this vein—that distinguishing between various accounts of risk and insurance can help us to better understand the emergence of different welfare regimes and to reflect more clearly on what each was trying to achieve.

A. Dilemmas of responsibility

From the emergence of industrial insurance and workers' compensation through the first national pension and health-insurance schemes, the vast majority of early social legislation was addressed to the problem of accidents, broadly construed. Policy debates from this period reveal the extent to which the rubric of risk and insurance allowed for shifting the locus of responsibility for workplace injuries and other hazards of economic life from individuals to a collective or from one collective to another. They also show that subsumed within these shifts were competing moral claims: about the power of human agency versus impersonal environmental forces; about the priority of individual liberty versus a common good; and about the salience of moral desert versus equality based on universal human needs.

In all of the countries that will be considered here, work-injury compensation laws tended to herald the movement toward more generalized social insurance. Such laws, which began to shift the

financial burden of workplace accidents from workers to their employers or the general public, hinged in part on a change in the understanding of causation and thus who should be responsible for industrial injuries. This development found different expressions across countries and times, but on the whole the late-19th-century trend away from employee responsibility is strikingly common in places as politically diverse as Germany, France, England, and even the United States. Although they did not always directly precipitate or pave the way for centralized social insurance, these parallel developments all reveal how the concept of risk, understood in statistical terms, both reflected and further enabled a significant shift in the understanding of human agency and moral responsibility.

In Germany, for example, a presumption of employer liability for work accidents was established in 1871, and was followed in 1885 by a law mandating employer insurance, the second of Bismarck's famous social insurance triad.²⁹⁹ In France, a German-inspired proposal failed in 1880, but others based on the principle of "occupational risk" succeeded in establishing compulsory insurance and statutory compensation for certain categories of workers.³⁰⁰ In England, an 1897 worker's compensation act allowed those suffering injuries from specific employments to recover compensation without proof of employer negligence, and another in 1906 extended this coverage to industrial diseases and most previously excluded occupations.³⁰¹ In the United States, the judicial attack on orthodox causation in tort law, the locus of worker-employer disputes, was encapsulated in

²⁹⁹ Deltev Zöllner, "Germany," in *The Evolution of Social Insurance*, 25-6.

³⁰⁰ Yves Saint-Jours, "France," 114.

³⁰¹ See A. I. Ogus, "Great Britain," in *The Evolution of Social Insurance*, 150-155. A very good analysis of the 1897 Act and its impact can be found in P. W. J. Bartrip and S. B. Burman, *The Wounded Soldiers of Industry: Industrial Compensation Policy 1833-1897* (Oxford: Clarendon Press, 1983), esp. 214-21. The authors argue that the law was not a working-class victory but rather a compromise solution that initially pleased no one. Its initial proponents were lawyers who sought to protect the sanctity of contract as they understood it in common law rather than extending the reach of tort-law liability. Employers for their part preferred it to the alternative of greater liability in tort. The ruling Conservative government supported it to gain working-class electoral support while avoiding a more interventionist approach that would have established national insurance or given the state a larger role in its administration. Nevertheless, Bartrip and Burman do accept that the act played an important role in the growth of public solutions to social problems, in recognizing that "the direct beneficiaries of a worker's labour owed him support in case of his injury in a work accident, irrespective of fault." *Ibid.*, 218.

1897 by Oliver Wendell Holmes, who argued that “the question of liability, if pressed far enough, is really the question of how far it is desirable that the public should insure the safety of those whose work it uses.”³⁰² Between 1911 and 1920, forty-two American states enacted workers’ compensation laws, followed by two more during the 1920s.³⁰³ In each country except France, full-fledged social insurance programs for sickness, old age, and unemployment were not long in following suit, all predicated on an understanding of certain eventualities as unavoidable or “social” risks, and therefore beyond the control of any given individual.

While the statistical understanding of accidents helped to reshape perceptions of individual responsibility, it did not imply any particular answer regarding to whom or what the burden of these events ought to be shifted. In France and subsequently Germany, the initial answer given was to the central government, with insurance understood as a direct and material contract between the individual worker and the state. In the wake of the revolution of 1848, France saw the establishment of its first national pension scheme, along with a contemporaneous law regarding mutual societies. In both cases, Napoleon III and his advisors sought, against the more liberal proposals of the legislature, to reinforce the role of the central government by directly subsidizing individual savings and recasting the mutuals as intermediaries between workers and the state.³⁰⁴

Bismarck, for his part, claimed to have learned from Napoleon the secret of using of social insurance to tie workers to the state, reportedly embracing the idea when he visited France in the

³⁰² Oliver Wendell Holmes, “The Path of the Law,” *Harvard Law Review* 10, no. 8 (1897): 467. See also Morton Horwitz, *The Transformation of American Law, 1870-1960* (New York and Oxford: Oxford University Press, 1992), 54-63.

³⁰³ As Theda Skocpol points out, these were the only reforms that succeeded out of the entire progressive-era social agenda. Skocpol, *Protecting Soldiers and Mothers*, 8.

³⁰⁴ Ewald, *L'État providence*, 210. Under these laws, the state would be granted responsibility for administering pensions, while the risks of sickness and accidents were to remain the province of the mutuals, under the protection and statistical tutelage of the government. Napoleon applied the same centralizing principles in his industrial accident insurance fund, established in 1868. Saint-Jours, “France,” 112.

mid-1850s and served as ambassador to Paris in 1861.³⁰⁵ He, too, assumed that responsibility for individual economic misfortune should be transferred directly from workers to the central government. In the case of accidents, he proposed mandatory, state-run employer insurance supported by a government subsidy. (The Reichstag, however, rejected both centralization and the subsidy in favor of preserving local insurance institutions.³⁰⁶) His mandatory health-insurance law, conceived together with the accident bill, also set out to replace voluntary institutions with centralized insurance provision, though he was ill on the day of its passage, and as a result it too preserved a greater role voluntarism than he would have liked.³⁰⁷ The third law in the triad, mandating disability and old-age pensions for all wage earners, was supposed to be financed through general taxation, thereby shifting to the social body as a whole the cost of workers' misfortune. In the end, however, considerations of financial practicability lead to the introduction of a contributory principle.³⁰⁸

In contrast to earlier forms of poor relief, these three laws all singled out specific groups for attention, primarily the employed and skilled core of the industrial working class. They also all set out to undermine the autonomy of working-class organizations, appealing directly to workers and establishing the state as their benefactor.³⁰⁹ Finally, all three operated on a principle of differentiated coverage, basing compensation on the worker's income, a feature that was related to their apparent moral purpose: to encourage stable employment and savings, and to give employees the idea that

³⁰⁵ Baldwin, *The Politics of Social Solidarity*, 102; Zöllner, "Germany," 12.

³⁰⁶ Zöllner, "Germany," 25.

³⁰⁷ E. P. Hennock, *The Origin of the Welfare State in England and Germany, 1850-1914: Social Policies Compared* (Cambridge: Cambridge University Press, 2007), 159-60 and Jürgen Tampke, "Bismarck's Social Legislation: A Genuine Breakthrough?" in *The Emergence of the Welfare State in Britain and Germany, 1850-1950*, ed. W. J. Mommsen (London: Croom Helm, 1981), 73.

³⁰⁸ Hennock, *The Origin of the Welfare State*, 183-87.

³⁰⁹ George Steinmetz, *Regulating the Social: The Welfare State and Local Politics in Imperial Germany* (Princeton, NJ: Princeton University Press, 1993), 124.

both they and their employers were investing in their future.³¹⁰ Social insurance would teach workers thrift and foresight—compelling them to save even when they were not so inclined—while establishing a contractual right to benefits, a perceived advance over more explicitly patronizing poor relief.³¹¹

The centralizing efforts of both Napoleon and Bismarck reflect the view that a direct insurance contract between workers and the state would create an identity of interests between the private and the public good. Individuals would get the security they need, while the state would secure their loyalty and with it a form of social order and control. As Bismarck explained in a parliamentary speech on behalf of his proposed pension law,

I lived in France long enough to realize that the attachment most Frenchmen feel towards their government...is essentially connected with the fact that most Frenchmen are in receipt of a state pension.... People there say, if the state comes to any harm, I'll lose my pension...so I have a vested interest in the state.³¹²

This was not, however, the only understanding of social insurance available to reformers of this period. The alternative, which regarded risks as properties of the statistical class rather than individuals, set out to transfer responsibility for accidents within and among risk groups, and to limit the extent of transfers to those who could not be considered equally responsible or vulnerable. Both visions rested on an understanding of accidents as beyond the scope of individual control, and both attempted to delineate some sphere of moral agency while affirming the overwhelming power

³¹⁰ Hennock, *The Origin of the Welfare State*, 195, 198; Steinmetz, *Regulating the Social*, 125.

³¹¹ See Baldwin, *The Politics of Social Solidarity*, 97 and Hennock, *The Origin of the Welfare State*, 198. Baldwin points out that in Germany the contractual nature of this right replaced the socialist argument for a right to aid from surplus economic value. Baldwin, *The Politics of Social Solidarity*, 71. Yet while this would appear to put social welfare on an entirely individualistic footing, the probabilistic argument could maintain both a contractual element and a claim from social utility or a social responsibility to compensate the worker.

³¹² Speech of June 18, 1889, in *Sten. Berichte des Reichstags*, VII, IV, vol. 3, 1831-36, 1834, quoted in Gerhard A. Ritter, *Social Welfare in Germany and Britain: Origins and Development*, trans. Kim Traynor (Warwickshire, UK: Berg Publishers, 1986), 35.

of social forces. Yet while the first account understood social insurance as a contract between the individual and the state, the second regarded voluntary associations and other self-defined risk groups as equally if not more significant to the purposes and workings of welfare. I submit that the latter account better characterizes many of the early social insurance programs enacted in Western Europe and the United States, particularly from the turn of the 20th century onward, than the liberal-statist vision that found expression in the policies of Napoleon III and Bismarck.

B. The targets of social policy

I have portrayed two theories of social insurance that emerged from the increasingly empirical character of probabilities in the 19th century: the late-classical or liberal statist and the frequentist or ethical collectivist. One of the main fault lines distinguishing the two accounts is the question of whether insurance is a reflection of responsibility to oneself and by extension the social order, or of solidarity with a statistical class, without which one's self-interest cannot be understood, let alone pursued. Bismarck's reforms, while designed for particular groups of workers, had as their ultimate target the individual, and set out to help him recognize his own vested interest in the state. The alternative vision, while also oriented toward risk-prone workers, targeted groups rather than individuals, and rested on a stronger sense of identity among their members. Although neither of these theories found wholesale expression in any actual welfare state, versions of the two did appear in many contemporary debates and found clear echoes in the policies that resulted.

In France, for example, after Napoleon III's pioneering reforms, it was an ethical-collectivist rather than liberal-statist account of insurance that dominated welfare provision. While an early innovator in social policy, France proved to be one of the slowest countries to develop an extensive centralized welfare state, thanks to the strength of its mutual societies and their liberal ethos of

voluntary aid.³¹³ While Bismarck had staunchly rejected voluntarism, social insurance in France remained deeply connected to its practice and ethical ideal. Often this limited what the state was able to accomplish: In 1910, for example, the French legislature tried to establish compulsory pensions for industrial and agricultural workers, but the measure was quickly undone by a court verdict declaring the obligation to contribute void. As late as 1925, an attempt to introduce limited unemployment insurance to prevent jobless mutualists from losing their coverage failed as well, as the proliferation and political clout of the mutual societies had created powerful vested interests against state provision.³¹⁴

The failure of these earlier laws, combined with the incomplete reach of the mutuals, did, however, serve as arguments for France's first major social insurance law, adopted in 1930.³¹⁵ Unlike earlier proposals, this law preserved a large role for mutual societies, allowing certain approved funds to provide health, maternity, and death coverage and others to carry disability and old-age schemes.³¹⁶ Mutualism in this context represented not only a powerful political force against state-run social insurance, but also an alternative vision of solidarity as more local and more sentimental, even more familial, than what could be achieved through an institution as vast and impersonal as the

³¹³ Saint-Jours, "France," 115.

³¹⁴ While the advance of the friendly societies had come to a halt in the final quarter of the 19th century, an 1898 law providing them with greater freedoms encouraged considerable growth in membership, mostly from the middle classes. The law was based on the work of historian and parliamentary specialist Hippolyte Maze, who unequivocally rejected state intervention in social relations, including the idea of compulsory insurance, and defended instead the mutualist principle as the best way to encourage social solidarity. See Abram De Swaan, *In the Care of the State: Health Care, Education, and Welfare in Europe and the USA in the Modern Era* (Cambridge: Polity Press, 1988), 198-201 and Gibaud, *De la mutualité à la sécurité sociale*, 49.

³¹⁵ In the years leading up to its passage, many mutualists had reluctantly concluded that some form of compulsory insurance was inevitable given the relatively limited scope of voluntary schemes, and they agreed to accept a compulsory plan in exchange for a central role in its administration. See Janet R. Horne, *A Social Laboratory for Modern France: The Musée Social and the Rise of the Welfare State* (Durham and London: Duke University Press, 2002), Kindle edition, 3906-3912.

³¹⁶ Saint-Jours, "France," 120-23; De Swaan, *In the Care of the State*, 201-2.

state.³¹⁷ This was class-based risk pooling and solidarity, with all its attendant advantages and limitations, on a local and voluntary scale.

The strength of mutualism in France also helps to explain why the most influential argument for social insurance at the end of the 19th century and beyond was one that effectively tried to transpose the friendly societies' ethos of mutual responsibility onto society as a whole. This was not the liberal-statist vision of Quetelet and Napoleon III, but rather the ethical-collectivist one of politician Léon Bourgeois. Bourgeois, who supported voluntarism, was responsible for forging the compromise position that persuaded mutualists to accept compulsory state insurance.³¹⁸ His account of solidarity, explained in his 1896 book of the same name, is explicitly reminiscent the probabilistic account of aleatory contracts.³¹⁹ Yet rather than agreeing *ex ante* to divide their common pool in accordance with each individual's expectation, as the original probabilistic doctrine had required, the participants in Bourgeois's society do so *ex post*, in accordance with what they have already received and the interdependence they experience as parts of the social whole. Peter Baldwin calls this "a Gallic version of the liberal idealist theories of citizenship spoken for across the Channel by T. H. Green, Alfred Marshall, R. H. Tawney and others that would later be given

³¹⁷ During the mid-19th-century debates over the national pension scheme, legislative reporter Benoist D'Azy argued against national insurance on the ground that mutual "societies are above all destined to propagate the spirit of family; these are real families, in which each one must be known, such that common respect is maintained in a spirit of reciprocity and duty." An insurance scheme based on too expansive a pool, as in compulsory state insurance, "will lose a large part of these advantages." Benoist D'Azy, Rapport au nom de la Commission chargée d'examiner les proposition de MM Dufournel et Lestiboudois..., Assemblée législative, October 6, 1849, *Annales*, 288, quoted in Ewald, *L'État providence*, 211. The continued strength of French mutualism may be attributable in part to the persistence of localism in French politics more generally: According to Timothy Smith, it was only between roughly 1880 and 1914 that Frenchmen began to shift their allegiance from their localities to the nation as a whole. See Timothy B. Smith, *Creating the Welfare State in France, 1880-1940* (Montreal and Kingston: McGill-Queen's University Press, 2003), 19.

³¹⁸ Horne, *A Social Laboratory for Modern France*, 3794.

³¹⁹ See Leon Bourgeois, *Solidarité* (Paris: A. Colin: 1896), 113-14: "For the determination of the rights and obligations of each in the solidaristic association that exists between men, for the calculation of profits and costs to be divided among all, one must be in possession of a coefficient common to all, a value of equal rights for all. Among innumerable elements of the calculation, drawn from the natural inequalities that separate and differentiate men, it is necessary always, to determine the equitable situation for each, to take into consideration this value and to admit it as equal for all."

concrete substance in Beveridge's reforms."³²⁰ I wish to emphasize that it also bears a strong resemblance to the frequentist account of the priority of the series, with the reference class defined as all parties to the social contract. The type of sympathetic solidarity this supports, based on an enlarged fellow feeling or analogy between the other's fate and one's own, is strongly reminiscent of the frequentist argument as well.

Bourgeois's account of mutualism sought to reconcile the solidaristic spirit of the mutuals with the centralizing tendencies of the age. Yet the French were apparently not yet disposed to regard the entire nation as a single risk pool, and the alternative—social insurance organized on the basis of classes—still struck many reformers as objectionable.³²¹ It was consequently only after World War II, in 1946, that coverage of perceived social risks was fully centralized and systematized in the form of social security, enlarging the role of the state and sounding the death knell of the mutualist movement.³²² Until then, French welfare provision remained largely the province of self-defined collectives, showing that a change in understandings of responsibility need not lead to centralized or compulsory welfare provision.³²³ Rather, on a frequentist view, the state is necessary mainly insofar as it facilitates risk pooling by or on behalf of self-defined groups that cannot do so effectively on their own.

In countries where an ethical-collectivist understanding of risk did lead to state-run welfare schemes, reformers and the interest groups they represented had fewer qualms about utilizing social

³²⁰ Baldwin, *The Politics of Social Solidarity*, 35.

³²¹ French politicians at this time were more concerned than their contemporaries in England and the United States about the effects of limited risk pooling in public insurance, which would privilege certain categories of beneficiary and potentially divide rather than unite the nation as a whole. Douglas Ashford, *The Emergence of the Welfare States* (Oxford: Basil Blackwell, 1986), 153-54.

³²² Gibaud, *De la mutualité a la sécurité sociale*, 120.

³²³ In France, the notion of "occupational risk," which affirmed that laborers were not responsible for workplace accidents, was well enshrined in law by the late-19th century, even as lawmakers rejected the idea of compulsory insurance. See Saint-Jours, "France," 114.

insurance for the benefit of particular classes. In Denmark, for example, the first social insurance scheme—a universal, tax-financed pension plan enacted in 1891—was established as a direct result of class-based risk pooling on the frequentist model. The law resulted from a protracted political struggle in which the agrarian bourgeoisie, unable to enact desired fiscal reforms, opted to use social policy for its own economic ends.³²⁴ Tax-financed pensions promised to shift the burden of poor relief from local land taxes to the central authority's indirect consumption levies, and therefore result in lower costs to the farmers.³²⁵ Agrarian employers also preferred a universal approach in order to appeal to a heterogeneous work force, since they were desperately in need of more manpower and relied on a variety of types of labor.³²⁶ The pension law was thus one of the primary concessions they extracted for themselves upon joining forces with moderate Conservatives to resolve the country's constitutional crisis. A similar political calculation, but very different outcome, characterized the second piece of Danish welfare legislation, a limited employers' liability law for work accidents passed in 1894. In this case, farmers rejected universal tax financing because it would have required them to share the burden of hazardous industrial practices in which they did not participate and from which they received no benefit. The resulting law consequently reflected narrow, industry-specific risk pooling rather than the broad coverage of the pension scheme.

Unlike in Germany, then, where social insurance set out to show workers that the state is a benevolent “institution...serving their needs and interests,” early Danish social legislation was a direct product of class mobilization, and in particular the desire of farmers to avoid redistribution to urban industrial laborers, who faced their own specific set of risks.³²⁷ The case of Denmark

³²⁴ Baldwin, *The Politics of Social Solidarity*, 65.

³²⁵ *Ibid.*, 74-5.

³²⁶ *Ibid.*, 81.

³²⁷ Preamble to the first draft of the accident insurance law of 1881, *Sammlung sämtlicher Drucksachen des Reichstags*, IV. Legislaturperiode IV, Session 1881, vol. 1, No. 41, 17, quoted in Ritter, *Social Welfare in Germany and Britain*, 83.

therefore reflects precisely the kind of class-based welfare that a frequentist account of social insurance would predict: an effort on the part of self-defined equals to pool their own risks in the name of collective security, while limiting redistribution to those considered outside of their reference group.³²⁸

C. The flexible actuarialism of early social insurance

Over the first few decades of the 20th century, this flexible, class-based approach to social policy became increasingly common. The political clout of self-defined groups grew, thanks to the spread of democracy and the organization of workers themselves, and new risks and population groups were increasingly incorporated into social insurance schemes. As a result, many welfare states expanded over this period from offering accident compensation and old-age pensions for limited groups of workers, to covering additional economic hazards and wider swaths of the working population. These new programs defined their aims pragmatically, expanding or contracting recipient groups and identifying new risks to serve diverse political needs. This pragmatic flexibility echoes the conclusions of frequentism, which acknowledged the reference-class problem and the relativity it injected into risk calculations, and maintained that only experience can reveal whether a risk is randomly distributed within a given population. Both frequentism and early welfare programs embraced an experimental approach to risk management that rested on certain assumptions about equality but remained open to reevaluating them as needed or desired.³²⁹

³²⁸ In Germany, the regime could afford to ignore the bourgeoisie because the latter had failed in 1848 to gain access to state power and remained marginal thereafter. See De Swaan, *In the Care of the State*, 189.

³²⁹ This account of early welfare's flexible actuarialism both echoes and departs from Karl Polanyi's influential account of the rise of economic interventionism in the later 19th century. Polanyi argues that the replacement of a liberal principle in economic life with a collectivist one did not arise from a "change either in the type of interests involved, or in the tendency of the opinions brought to bear on the matter," but rather from "the evolving conditions under which the problem arose and a solution was sought." Karl Polanyi, *The Great Transformation: The Political and Economic Origins of Our Time* (Boston, MA: Beacon Press, 2001), 153. My discussion accords with Polanyi's in emphasizing the pragmatic

The evolution of British social insurance offers a powerful illustration of the logic just described. England's early welfare state was inspired directly by the German example, and as a result affirmed certain liberal-statist principles.³³⁰ Compulsory pensions, for instance, enacted in 1908, provided flat-rate payments financed by general taxes and paid to every citizen over the age of seventy with an income of less than a certain amount per year.³³¹ In principle, rather than targeting any particular class of workers or citizens, pensions were designed to be universal in reach and distributed on the basis of need. In practice, however, considerations of cost meant that coverage was made contingent on good character, requiring applicants to prove their mental health, their prior avoidance of poor relief, and their habitual employment in the trade of their choice.³³² The choice of tax financing as opposed to contributions—which, we recall, had been Bismarck's own preference for pensions—was based partly on the view that the destitute could not be expected to save for themselves, though it was also a capitulation to friendly societies, who feared that a contributory system would deflect working-class savings and reduce their own ranks.³³³

nature of the early collectivist approach to welfare, but diverges insofar as I see this pragmatism as also reflecting philosophical principle.

³³⁰ On the British adoption of German social insurance, see E. P. Hennock, "The Origins of British National Insurance and the German Precedent, 1880-1914," in *The Emergence of the Welfare State in Britain and Germany*, 84-106.

³³¹ As in the other countries we are considering, in England several laws on employer liability and workmen's compensation had gradually extended laborers' rights of recovery for industrial accidents between 1880 and 1906. In 1905, a legislative resolution allowing local authorities to give food to hungry schoolchildren officially established the principle of state concern for the lives of adult laborers (via their offspring), and was followed a few years later by the institution of school medical inspections. See Bentley Gilbert, *The Evolution of National Insurance in Great Britain: The Origins of the Welfare State* (London: Michael Joseph, 1966), ch. 3. Despite these earlier developments, however, many historians seem to agree that the new era of social reform did not begin in England until the introduction of old-age pensions in 1908.

³³² *Ibid.*, 222.

³³³ Bentley Gilbert, "The Decay of Nineteenth-Century Provident Institutions and the Coming of Old Age Pensions in Great Britain," *The Economic History Review* 17, no. 3 (1965); cf. Baldwin, *The Politics of Social Solidarity*, 99-100; Harris, *Origins of the British Welfare State*, 164; and Hugh Hecló, *Modern Social Politics in Britain and Sweden: From Relief to Income Maintenance* (New Haven and London: Yale University Press, 1974), 175.

While the pension law had echoes of Bismarck, then, unemployment insurance, enacted in 1911 together with health and invalidity provisions, rested on a very different justification.³³⁴ Until this time, unemployment had been considered uninsurable, a cumulative risk that could not be effectively predicted.³³⁵ Yet the economic depressions of 1879 and 1908 had led the British public to conclude that unemployment was a product of economic forces far more powerful than particular workers.³³⁶ Liberal reformers in the British government consequently set out to do for workers collectively what they arguably could not do for themselves. As Winston Churchill, then president of the Board of Trade in Asquith's government, explained in a 1909 speech, insurance would cover "trades in which unemployment is not only high, but chronic...and in which, wherever and howsoever it occurs, it takes the form not of short time or of any of those devices for spreading wages and equalizing or averaging risks, but of a total, absolute, periodical discharge of a certain proportion of the workers."³³⁷ In such cases, individual foresight or prudence is irrelevant, because no voluntary means could possibly accomplish the necessary end.

In keeping with this view, unemployment insurance assumed a lateral model of redistribution, with contributions from those currently employed in a given trade financing the claims of those contemporaneously out of work.³³⁸ The rationale for this structure was given expression by William Beveridge, one of the scheme's main proponents at the time. Against the

³³⁴ Hecló also emphasizes the centrality of unemployment insurance to the development of welfare in both Great Britain and Sweden, although he accounts for it in terms of policymakers' failure to "break the conceptual hold of classical economics" rather than as a reflection of an alternative approach to risk-management, as is my contention here. Hecló, *Modern Social Politics*, 91.

³³⁵ Bismarck had avoided the subject, and German unemployment legislation would have to wait until 1927. De Swaan, *In the Care of the State*, 196.

³³⁶ Ibid.

³³⁷ Winston S. Churchill, "Labour Exchanges and Unemployment Insurance," House of Commons, May 19, 1909, in *Liberalism and the Social Problem* (New York: Haskell House Publishers, 1973), 267.

³³⁸ Gilbert, *The Evolution of National Insurance in Great Britain*, 281.

view of Sydney and Beatrice Webb, who advocated a reformative approach to the unemployed, Beveridge emphasized that many workers face an unavoidable difficulty of “averaging wages over good times and bad.” He further justified the expansive reference class of all workers on the ground of their fundamental interdependence: “The regular workman must admit a certain solidarity...with the irregular workman, since without the latter the industry by which the former lives could not be carried on.”³³⁹ The case for solidarity depended in this instance on a kind of assumed probabilistic equality among workers, and on the sympathy it generates for those who are roughly similar yet face hardship through no fault of their own. In a sense, this was the collective solidarity of the mutual societies applied, thanks to the coercive powers of the state, on a more inclusive basis. Where mutual societies had failed to cover the most difficult risks, the state could mandate workers to share burdens not only within but also beyond their actuarial class.

In the United States, as in nearly every industrial democracy, the issue of work accidents played a foundational role in the gradual emergence of the welfare state, providing the paradigmatic example of an undeserved misfortune in need of redress. As elsewhere, this problem led to the emergence of various cooperative workers’ associations that provided members with rudimentary insurance-type relief. These were successful in cultivating a strong sense of mutuality and avoiding the challenges inherent to disability insurance, in particular the self-selection of high-risk candidates and the potential for fraud.³⁴⁰ By the 1890s, cooperatives were also the leading source of life insurance in the United States, reporting a substantially greater sum of policies in force than commercial mutual and stock companies combined.³⁴¹

³³⁹ William Beveridge, “Unemployment Insurance, Objections and Answers,” quoted in José Harris, *William Beveridge: A Biography*, revised paperback ed. (Oxford: Oxford University Press, 1997), 181.

³⁴⁰ John Fabian Witt, “Toward a New History of American Accident Law: Classical Tort Law and the Cooperative First-Party Insurance Movement,” *Harvard Law Review* 114, no. 3 (2001): 781.

³⁴¹ *Ibid.*, 778.

In contrast to what transpired in both England and France, however, the American welfare state did not retain an important role for such voluntary insurance societies in its eventual administration. This was in part a product of the cooperatives' own intransigent commitment to voluntarism and their rejection of any state role in providing insurance.³⁴² It was also the result of an unfriendly legal climate and a massive influx of poor immigrants, both of which meant that the cooperatives had lost much of their comparative advantage by the time reformers began to push for broad social insurance around 1910.³⁴³

The limited influence of friendly societies, both in principle and as a political movement, may help to explain why the model of mutualism and its strong solidarity played a more muted role in the development of American social welfare than was the case in England and France. Since the cooperative societies were not perceived as a viable alternative to social insurance at the time the movement for the latter took off, welfare advocates did not take its approach of limited and self-policed risk-pooling seriously. For example, Isaac Rubinow, the most prominent early 20th century advocate of social insurance in the United States, saw compulsory state insurance as different in kind from both private insurance, with its profit motive and reliance on individual savings, and mutual societies, which on his account operated on the model of charity rather than right.³⁴⁴ Only a mandatory, publicly administered insurance scheme could “substitute a social guarantee against the results of emergencies and accidents for the purely personal responsibility which is still the rule in

³⁴² According to Witt, “many cooperative association voices evinced a deeply rooted commitment to insurance as a voluntarist mechanism for accommodating the individual to the dilemmas of industrial life. From this perspective, insurance allowed individuals to solve the social problems that otherwise threatened to result in expansion of the state. The effective use of insurance could thus counter calls for new governmental functions.” *Ibid.*, 811.

³⁴³ *Ibid.*, 828-32, 837.

³⁴⁴ See I. M. Rubinow, “Labor Insurance,” *Journal of Political Economy* 12, no. 3 (1904): 362-63. This view of the friendlies as offering charity is historically inaccurate, at least as a matter of the self-understanding of society members themselves. Cordery argues that as early as 1800, friendlies were widespread precisely because “members could claim [insurance] as a right rather than request [it] as a gift.” Cordery, *British Friendly Societies*, 18.

many countries.”³⁴⁵ Unlike in France, then, where voluntary mutual assurance and social insurance had long been seen as resting on similar philosophical foundations, in the United States there was no such tradition. The idea of risk pooling on a social scale appeared far more foreign and threatening to Americans’ predominantly liberal sensibility.³⁴⁶

As a result, of the countries under consideration here, the United States seems least to illustrate the influence of frequentist ideas in the emergence of social policy. Early American welfare did focus on specific risks—injury, sickness, and unemployment—that were understood as systemic in nature and therefore beyond individual control. But the influence of risk-based classes of workers, whether organized as political movements or in mutual insurance societies, appears to have been far more muted in the American context than in Denmark, England, and even France.

There was, however, one self-defined risk class that was particularly well organized and influential in the United States: that of employers.³⁴⁷ In a sense, then, the eventual shape of the American welfare state does reflect the aleatory collectivism associated with a frequentist account of probabilities, but as shaped to a noteworthy extent by the risks of private enterprise. For instance, American workers’ compensation statutes, enacted in 45 states between 1911 and 1920, were limited in their coverage and tended to reflect the demands of the private sector, which welcomed them as reducing the administrative burden of accident litigation and lowering their own insurance

³⁴⁵ I. M. Rubinow, “Compulsory State Insurance of Workingmen,” *The Annals of the American Academy of Political and Social Science* 24 (1904): 47.

³⁴⁶ See Michelle Landis Dauber, *The Sympathetic State: Disaster Relief and the Origins of the American Welfare State* (Chicago: University of Chicago Press, 2013), Kindle edition, 379 and Walter I. Trattner, *From Poor Law to Welfare State: A History of Social Welfare in America*, 6th ed. (New York: Free Press, 1999), 278-79.

³⁴⁷ On the powerful role played by private employers in the passage of the Social Security Act, see Jacob S. Hacker, *The Divided Welfare State: The Battle over Public and Private Benefits in the United States* (New York: Cambridge University Press, 2002), Kindle edition, 97-104.

premiums.³⁴⁸ Even after the New Deal's expansion of federal welfare provision, the American welfare state continued to reflect the contractual orientation of private employers. As Jacob Hacker points out, the Social Security Act succeeded in part because private organizations had provided only spotty and insufficient pension coverage, and as a result employers had little vested interest in the continuation of the existing system.³⁴⁹ Although many business leaders did oppose the act, they retained an organized influence over its workings, particularly in securing the resemblance of public welfare to private arrangements.³⁵⁰ Moreover, companies soon learned to use Social Security to their advantage, combining corporate pension plans with federal benefits to attract and retain better-paid employees.³⁵¹ The American case is perhaps a reminder, then, that class-based perceptions of risk need not confine themselves to workers, and that their distributive implications are not always or inherently egalitarian.³⁵²

With some exceptions, the early welfare programs I have been considering originated as a response to the plight of workers, in particular the challenges inherent in their reliance on continued

³⁴⁸ Like other Progressive Era minimum standards and social insurance laws passed by the states, they relied on employers to provide for their employees rather than establishing provision directly from the state. See Edward Berkowitz and Kim McQuaid, *Creating the Welfare State: The Political Economy of Twentieth-Century Reform*, revised ed. (Lawrence, KS: University Press of Kansas, 1992), 43-49.

³⁴⁹ Hacker, *The Divided Welfare State*, 106-7.

³⁵⁰ Berkowitz and McQuaid, *Creating the Welfare State*, 126, 156. Business leaders did not succeed in achieving all of their aims, however. For example, the Clark Amendment to the Social Security Act would have allowed employers and employees who already had pension plans to opt out of the federal law. Planners objected on the grounds that this would leave the federal government with the worst risks to insure and saddle the public sector with prohibitive costs. Although the measure passed both houses of Congress, it was tabled by Roosevelt and ultimately abandoned. See *ibid.*, 125.

³⁵¹ Hacker, *The Divided Welfare State*, 104.

³⁵² For a detailed account of the role of interest-group preferences, including and particularly business interests, in the emergence of welfare benefits in the United States, see Frank R. Dobbin, "The Origins of Private Social Insurance: Public Policy and Fringe Benefits in America, 1920-1950," *American Journal of Sociology* 97, no. 5 (1992): 1416-1450.

salaried work for survival. As the legal scholar A. V. Dicey put it in 1917, explaining the rise of collectivism in the 19th century, “The sale of labour is felt to be unlike the sale of goods. A shopkeeper can keep back his wares until the market rises, whilst a factory hand, if he refuses low wages, runs the risk of pauperism or of starvation.”³⁵³ By either compelling workers to save for their own futures or pooling risks and redistributing their costs among designated equals, social insurance could ensure that faultless interruptions in a worker’s earnings would not result in economic devastation. Put another way, social insurance implicitly answered the questions of who would be allowed to opt out of the labor market and who would in turn shoulder the costs of their doing so. In many of the early programs, the answer to the first question turned out to be workers struck by risks considered systemic and randomly distributed, while the answer to the second question was primarily the workers themselves, either as individuals or as a class, or both.³⁵⁴

It is certainly not accidental that the concept of risk, which emerged as a commercial tool to help merchants limit their losses, was appropriated to save industrial capitalism by helping workers limit the downside of their participation in the labor market. Of course, the understanding of risk also underwent a serious modification over the course of the 19th century, coming with the help of probabilistic insurance to embody an ideal of interpersonal redistribution that could support a public, collective response to uncertainty. It was very likely this dual character of the concept—its close relation to the logic of contractual exchange and its simultaneous redistributive potential—that made it so appealing as a framework for social legislation. This same dual character in turn suggests

³⁵³ A. V. Dicey, *Lectures on the Relation between Law and Public Opinion in England during the 19th Century*, ed. Richard VandeWetering (Indianapolis: Liberty Fund, 1917), 188, <http://oll.libertyfund.org/titles/2119>.

³⁵⁴ In emphasizing the class-oriented character of early social legislation I depart somewhat from Pierre Rosanvallon’s characterization of the welfare state as hinging on ignorance of relevant actuarial differences. While this may have been true of the post-World War II solidaristic welfare state, it was not quite true of the earliest social insurance programs. See Rosanvallon, *La nouvelle question sociale*, ch. 1.

that efficiency-based and altruistic arguments for social welfare may in fact have a common root, in the pragmatic solidarity of frequentism and its flexible account of probabilistic equality.

Chapter Four: The Egalitarian Welfare State and the Ambiguities of Insurance

In the last chapter, I argued that frequentist probability's rejection of the Laplacean paradigm, and in particular its insistence that probabilities can be assigned exclusively to statistical classes and not to individual events, was associated with a novel defense of insurance. The frequentist account was collectivist rather than liberal, and found practical echoes in the group-based welfare policies enacted in the decades just before and after the turn of the 20th century. The appeal of such class-based risk-pooling was that it could be simultaneously a prudent means of self-protection in a market economy and a kind of circumscribed altruism—a commitment to band together with one's "own" in times both good and bad, just like the fraternal societies out of which the welfare state grew and which it eventually replaced as workers' main source of economic security.

The main difference between the friendlies and early social insurance on the account I am proposing is that while the former relied heavily on communal sentiment to supplement their implicit actuarialism, the latter turned to the state and its powers of coercion to enforce a more explicit account of probabilistic equality. The frequentist case for insurance thus combined an individual rationale—an imagined similarity or interchangeability with those one deems relevantly similar—with a broadly communal or social result. Thanks to this dual appeal, the insurance principle eventually superseded need and desert, previously the main justifications for distributive claims, as the primary rationale for public welfare provision.

The years between the two world wars and after 1945 further established the insurance principle as the core of social welfare policy in the advanced European democracies and the United States. Many of the expansive welfare programs enacted during this era invoked the idea of broad risk pooling to promote social integration, self-respect, and basic security for all. At the same time, these programs inevitably, and often intentionally, departed so substantially from any view of

actuarial fairness that in some respects they ceased to resemble insurance at all. As a result, while the rhetoric of risk remained essential to political support for welfare policies, particularly among the increasingly influential middle classes, the normative aims of welfare extended well beyond what empirical probabilistic justice could achieve. A number of mid-century social theorists consequently concluded that welfare programs ought to abandon or transcend the logic of insurance altogether.

At approximately the same time but for different reasons, developments in probability theory and economics also cast doubt on the earlier vision of social insurance as an expression of solidarity among aleatory peers. While the frequentist account of probability had supported risk-pooling efforts by particular groups and by technocrats acting on their behalf, the subjective theory of risk that rose to prominence in the 20th century offers no clear justification for probabilistic collective action. On the contrary, by understanding mathematical likelihoods as above all a tool of individual instrumental reason, the subjective account downplayed concerns about interpersonal equity and tended to reduce all reasoning to a form of wager. I do not intend to suggest that this shift was responsible for growing concerns about the insurance model in mid-century welfare practice or theory. I do, however, maintain that both trends found expression in the most influential work of 20th-century political thought, John Rawls's *Theory of Justice*. Rawls's account of distributive justice can be read in the context of this political and intellectual history as an effort to rescue social insurance from itself. By reading Rawls's work in this way, I hope to shed light not only on the development of 20th-century political theory but also, and more deeply, on some of the challenges inherent to the liberal welfare state.

I. The egalitarian welfare state emerges

A. The centrality of insurance to post-war egalitarian welfare

Many histories of the welfare state attribute the emergence of modern social policy to the years immediately following World War II, when, according to the standard narrative, recent memory of the war's indiscriminate devastation paved the way for sweeping protective legislation.³⁵⁵ Indeed, many of the programs that define the era of universal centralized welfare, or what we might call the egalitarian welfare state, date from 1945 or thereafter. In England, three major social policy statutes—the Family Allowances Act of 1945, the National Insurance and National Health Service Acts of 1946, and the National Assistance Act of 1948—were enacted immediately after the war based on the ideas laid out by William Beveridge in his 1942 *Social Insurance and Allied Services*, popularly known as the Beveridge Report.³⁵⁶ Beveridge and the Labour government that instituted his proposals thus established the principle of universal contributory coverage, providing flat-rate benefits to all as a matter of right, at the heart of the British welfare state and as an aspiration for social legislation worldwide. Universal flat-rate pensions were passed with broad political support in Sweden in 1946 and in Denmark in 1956.³⁵⁷ The post-war period also saw a variety of efforts to enact similar programs in France and Germany, though these came to fruition only in the 1960s and

³⁵⁵ The authoritative articulation of this view came from Richard Titmuss, "War and Social Policy" in his *Essays on the Welfare State*, 3rd ed. (London: George Allen & Unwin, 1966), 75-87. See also Asa Briggs, "The Welfare State in Historical Perspective," *European Journal of Sociology* 2, no. 2 (1961): 223.

³⁵⁶ Sir William Beveridge, *Social Insurance and Allied Services*, American ed. (New York: Macmillan, 1942). For a full account of Beveridge's report and its implementation see Harris, *William Beveridge*, 378-451.

³⁵⁷ Peter Baldwin, *The Politics of Social Solidarity*, 140-43 and 152-53. A flat-rate pension was also created in the Netherlands in the immediate post-war period, while in Canada, as in Sweden and Britain, means-testing was abandoned as part of the push for a more universal, Beveridge-style old-age assistance system. See Cutler and Johnson, "The Birth and Growth of the Social Insurance State," 98-9.

70s, around the same time that several national welfare programs were created in the United States as well.³⁵⁸

The adoption of expansive, Beveridge-style programs marks a change in the aims of social policy from helping particular groups of risk-prone workers to providing a basic level of security for everyone.³⁵⁹ Both the insurance programs instantiated at this time, which were designed primarily to provide income support, and the social services introduced to supplement them were organized along these lines, offered to all at uniform standards regardless of status or earnings.³⁶⁰ As a result, both contemporary and recent commentators have stressed a shift during this period from what one early study called the “cannon...of collective risk, that is the standardization of contributions and benefits” within statistical classes, to the “cannon of a subsistence minimum guaranteed through social security” with the help of state subsidies.³⁶¹

While this was undoubtedly a noteworthy change, the continuities in social welfare between the pre- and post-war periods are equally significant.³⁶² As we saw in previous chapters, the idea of insurance was at the center of social policy from its inception, and only grew in prominence as the 20th century progressed. As a result, more recent efforts to distinguish various eras of welfare

³⁵⁸ Baldwin, *The Politics of Social Solidarity*, 279-86 and Hacker, *The Divided Welfare State*, 21.

³⁵⁹ According to Peter Baldwin, prior to Beveridge social insurance across Europe had not fulfilled its “solidaristic potential,” because it spread risks only within socioeconomically homogeneous groups. Beveridge’s ambitions to cover all people against all risks were totalizing by contrast. Peter Baldwin, “Beveridge in the *Longue Duree*,” in *Beveridge and Social Security: An International Retrospective*, ed. John Hills, John Ditch, and Howard Glennerster (Oxford and New York: Clarendon Press, 1994), 42-5.

³⁶⁰ See Robert E. Goodin, “The End of the Welfare State?” in *The Cambridge History of Twentieth Century Political Thought*, 203-4.

³⁶¹ Alan T. Peacock, *The Economics of National Insurance* (London: William Hodge and Co. Ltd.: 1952), 41-42.

³⁶² Douglas Ashford has chronicled the ways in which the interwar period, while lacking in significant policy innovations, helped lay the institutional and political foundations for post-war social legislation. For example, during these years the advanced democracies of Europe and the United States experienced significant bureaucratic expansion, the ongoing substitution of national for local welfare provision, and the waning of mutual insurance societies. Douglas E. Ashford, *The Emergence of the Welfare States* (Oxford: Basil Blackwell, 1986), ch. 3.

provision—from the earlier social assistance state, to the insurance-dominated “social security” state, to the non-contributory “social welfare” state—tend to obscure the ongoing prominence of insurance rhetoric across time periods and programs.³⁶³ As social security expert Eveline Burns pointed out in her 1949 study of the American welfare system, it had long been common to refer to government programs as “insurance” even when their workings deviated from actuarial and even contributory principles.³⁶⁴ While the institution underwent a substantial evolution over time, covering “not only new risks but...new population groups” and eventually guaranteeing a universal basic minimum, it remained through all of these iterations widely understood as “social insurance,” a means to protect the vulnerable by pooling their resources and compensating the blameless sufferers among them.³⁶⁵

If we must divide the history of welfare provision into the periods before and after 1945, then, what distinguishes them is not the underlying insurance-based principle that distribution should shield against or compensate for undeserved hardship, a principle common to both eras, but rather the scope of the recipient group and the accounts of equality and solidarity thereby entailed. The paternalistic Bismarckian (originally Napoleonic) idea of tying workers to the state had already been supplanted in the 1910s and 20s by a more flexible egalitarian collectivism aimed at reciprocity

³⁶³ An example of such an attempt is Robert E. Goodin, *Reasons for Welfare: The Political Theory of the Welfare State* (Princeton: Princeton University Press, 1988), 3-4. For the same reason, I do not rely on the now widespread “Bismarckian” versus “Beveridgean” typology to distinguish between programs in which benefits are related to earnings and those that provide a flat-rate minimum to all. This distinction, while helpful in explaining the expansion of welfare coverage after the war, otherwise oversimplifies the political history of social insurance while minimizing the fact that both types of system sprang from common theoretical roots.

³⁶⁴ Workers were only rarely required to make contributions under American unemployment laws, for instance, and never under workers’ compensation, yet such programs were still widely understood as offering “insurance.” Eveline M. Burns, *The American Social Security System* (Boston: Houghton Mifflin, 1949), 29-31. Similarly, as Beatrice and Sydney Webb had admitted a few decades earlier, in discussing compulsory invalidity insurance, the issue of how precisely to finance such schemes “is, in reality, one as to what is the most equitable and most convenient incidence of taxation,” without regard for proportionality between contributions and benefits. Sidney and Beatrice Webb, *The Prevention of Destitution* (London: Longmans, Green and Co., 1911), 186.

³⁶⁵ Eveline M. Burns, “Social Insurance in Evolution,” *The American Economic Review* 34, no. 1 (1944): 201.

and mutual aid. Post-war social insurance simply carried the latter idea further, proposing a vision of universal equality based on common exposure to market-generated hazards and other forms of chance.

Why, then, did the insurance principle continue to hold such a grasp on the imaginations of social reformers and politicians even as the aims of welfare provision expanded beyond empirical probabilistic justice? Without intending a full explanation, I will suggest a few important factors. For one, path dependence likely played a significant role, in that early social policy decisions heavily influenced what was subsequently considered desirable and politically possible. In most of the countries under consideration here, the first public responses to the plight of industrial laborers relied heavily on insurance mechanisms and their providers, including commercial insurers (as in some workers' compensation schemes), mutual societies (where they helped administer public programs), and of course the state itself. These policies not only helped to legitimize insurance in principle, but they also created institutional, economic, and legal realities that then proved difficult to reverse.³⁶⁶ While the nature of such path dependence differed from country to country, and did not always lead to public rather than private schemes, it is almost certainly true that such initial decisions helped create the conditions for the continued preeminence of insurance as the favored mechanism of welfare provision.³⁶⁷

Another important factor was the advancement of mathematical statistics as a discipline, which developed new tools to interpret large-scale phenomena and acquired new prominence between the world wars. Probability theory assumed a more central role in both the natural sciences

³⁶⁶ An early version of this claim was made by Burns, who pointed out that existing social insurance programs exert a “dynamic influence,” broadening public perceptions of the aims of welfare and legitimating insurance as the means for achieving those aims. Burns, “Social Insurance in Evolution,” 199-211.

³⁶⁷ Hacker’s argument, for example, is that in the American case path dependence frequently encouraged private over public welfare provision. Hacker, *The Divided Welfare State*, 55-58.

and economic theory during the 1920s.³⁶⁸ Around the same time, the new technique of random sampling, which presumed that policymakers could make conclusions about an entire nation via surveys of representative groups, gave statisticians considerable importance and authority in increasingly centralized states.³⁶⁹ Statistics appears to have solidified its grasp on policymaking during the 1930s, thanks largely to the economic turmoil of that decade and the perceived need for centralized planning.³⁷⁰ Thus as a matter of sheer political visibility and clout, statistical thinking made inroads between the wars that in turn enabled its central role in the policy revolution that followed.

Finally, and perhaps above all, it was the conceptual and practical plasticity of insurance that allowed it to remain at the center of welfare policy both before and after 1945. I have been arguing that each major account of social insurance managed to reconcile in practice a pair of competing theoretical principles. In its post-war variety, insurance had the appeal of being egalitarian without being too drastically redistributive, and of acknowledging the salience of basic needs without relinquishing individual responsibility or desert.³⁷¹ The view that everyone is equally vulnerable to a given risk meant that welfare could cast off any stigma associated with the old Poor Law and means testing, a central concern of those who advocated universal benefits.³⁷² At the same time, the fact that the individual paid contributions out of his wages (or even taxes) meant that he had in some

³⁶⁸ Alain Desrosières, *The Politics of Large Numbers: A History of Statistical Reasoning*, trans. Camille Naish (Cambridge, MA: Harvard University Press, 1998), 164.

³⁶⁹ *Ibid.*, 204-9.

³⁷⁰ *Ibid.*, 165 and Briggs, “The Welfare State in Historical Perspective,” 229.

³⁷¹ For a discussion from this period that inadvertently shifts between the two distributive principles, see A. D. Watson, “The Principles Which Should Govern the Structure and Provisions of a Scheme of Unemployment Insurance” (Ottawa: 1948), esp. 14-15.

³⁷² See, e.g., Harris, *William Beveridge*, 392-93 and 417 and Eveline M. Burns, “Priorities for Public Welfare,” *Social Work* 3, no. 4 (1958): 38-9.

sense merited his coverage and any associated payout.³⁷³ Insurance thus promised to harmonize the claims of personal responsibility or merit with those of social equality or universal self-respect. This harmonizing character, as we will now see, proved to be at once a political necessity and a theoretical liability.

B. The amorphous appeal of insurance

The notion that social insurance could bridge the competing distributive principles of merit and need, individual responsibility and universal security, found expression in the work of its early advocates, not least of all Beveridge himself. The architect of the liberal egalitarian welfare state presented insurance as a kind of third way, the key to “maintaining individual freedom and responsibilities” while “giving security against all the main risks of economic life.” If Communism meant guaranteeing an income “at all times to everybody irrespective of his work and services,” and laissez-faire capitalism abandoned workers to the vagaries of the market, social insurance promised to protect the hardworking and responsible from disruptions due to “causes beyond their control.”³⁷⁴

In practice, Beveridgean welfare programs did this by accommodating a range of financing and benefit options—including general tax financing or differential contributions combined with

³⁷³ See also Burns, “Social Insurance in Evolution,” 207.

³⁷⁴ William Beveridge, “Insurance for All and Everything” (London: The Daily News Ltd., 1924), esp. 31. Harris asserts that during the interwar years Beveridge was utterly opposed to the kind of liberal-collectivist, mixed-economy principles that later became associated with his proposals, and instead believed that policymakers had to choose between a free market softened by a residual safety net or a totally regulated economy on the model of the Soviet Union. Yet even if his presentation of insurance here as a moderating force was more rhetorical than principled, he clearly changed his position during the war, coming to see social insurance as just one part of a far more ambitious economic program that would include minimum-wage legislation, the nationalization of land and housing, and the encouragement of full employment. Beveridge’s vision, which Harris somewhat inaptly characterizes as “civic republican,” was in fact deeply indebted to a middle-class ethos of economic independence and self-discipline. This in turn helps to explain his support for both voluntary and compulsory insurance as cultivating private moral virtues and encouraging the strength of family and communal life. José Harris, “Beveridge’s Social and Political Thought,” in *Beveridge and Social Security*, 28-35.

flat-rate payments—while still affirming, at least in the public perception, the idea that the individual received what he had paid for.³⁷⁵ Beveridge readily acknowledged that such insurance “preserves the contributory principle” only nominally, since the state, with its powers of compulsion and taxation, “is not under the necessity of varying the premium according to the risk.”³⁷⁶ Nevertheless, his writings explicitly and repeatedly affirmed the primacy of the insurance principle as enabling a system in which benefits are “earned” through contributions, means tests are avoided, and the state as benevolent planner orchestrates a comprehensive response to a variety of so-called social risks.³⁷⁷

Several decades earlier, Beveridge’s close associates Beatrice and Sydney Webb had explained the power of insurance in similar, if somewhat less enthusiastic, terms. “We are,” they wrote in their *The Prevention of Destitution*, “face to face with an obsession of the public mind in favour of insurance. This obsession is not likely to be removed by any demonstration that it depends on a confusion between voluntary and compulsory insurance, which have entirely different attributes, and lead to entirely different results.” Nor would ordinary citizens likely be persuaded that the latter “involves an extravagant expenditure of public funds on persons who would in any event have maintained themselves at the prescribed standard of civilized life.”³⁷⁸ Despite the Webbs’ lack of patience for the liberal prejudices of the English public, they understood the political appeal of the insurance principle and were ready to use it as a means to more comprehensive ends. The “more universal and more compulsory” the insurance scheme, “the more quickly and the more certainly” it will give rise to a “Policy of Prevention,” since there will be greater public incentive to prevent self-

³⁷⁵ Britain’s National Health Service, for example, was from the outset to be financed by general taxation rather than individual contributions. A. I. Ogus, “Britain,” 192.

³⁷⁶ Beveridge, *Social Insurance and Allied Services*, 12-13.

³⁷⁷ *Ibid.*, 193-94. It is worth noting here that the abolition of means tests also had a liberal rationale, in that it was intended to encourage private self-help. This fact helps in part to explain the appeal of universal post-war reform to parties not typically associated with it. Baldwin, *The Politics of Social Solidarity*, 116.

³⁷⁸ Sidney and Beatrice Webb, *The Prevention of Destitution*, 213.

indulgence and other irresponsible behaviors on the part of the insured.³⁷⁹ Until then, social insurance policies relying on the public's irrational enthusiasm provided what seemed to be a path toward greater social cohesion using familiar terminology and relatively flexible means.

The Webbs' preventative approach never carried the day, and Beveridge's own writings emphasized the moderate character of insurance rather than laying out a vision that would eventually supersede it.³⁸⁰ Yet the question of whether universal social insurance was an end or a means to more comprehensive goals remained a central one among reformers and social theorists in the decades after 1945. If the virtue of the insurance principle is its responsiveness to both merit and need, it will encourage very different distributive outcomes than a system that takes account only or primarily of the latter.³⁸¹ The Beveridgean welfare state thus begged the question of whether social insurance could be a means to its own abolition or whether, given the nature of its appeal, it would forever remain tied to class-based distributive politics.

The preeminent welfare theorist of this period, Richard Titmuss, in describing his own support for universal insurance programs after the war, explained that he had understood these measures as "connected with the demand for one society," characterized by universal self-respect and social integration, rather than with the redistribution of wealth for its own sake.³⁸² Titmuss was

³⁷⁹ Ibid., 214; cf. 185-92. Although the Webbs were opposed to unconditional relief, they were in favor of universality as a means to break up the Poor Laws, and to this end promoted national tax financing. See Gilbert, *The Evolution of National Insurance in Great Britain*, 214-16 and 271. Their thinking about the organization of social insurance schemes directly and significantly influenced Beveridge, who remained faithful to their earlier framework even as they became increasingly despondent about the possibility of reforming the Poor Laws. Harris, *William Beveridge*, 354-55.

³⁸⁰ On the Webbs and the relationship between prevention and universalism, see Richard Titmuss, "Welfare State and Welfare Society," in *The Philosophy of Welfare: Selected Writings of Richard M. Titmuss* (London: Allen & Unwen, 1987), 147-48. On the limited influence of Fabian ideas on the shape of the British welfare state, see E. J. Hobsbaum, *Labouring Men: Studies in the History of Labour* (London: Weidenfeld and Nicolson, 1964), 252.

³⁸¹ Occupational welfare, for example, distributes on the basis of income or status in addition to need, and often with regressive effects. Richard Titmuss was among the first to call attention to this problem, which has more recently become a central concern of welfare studies. See his "The Role of Redistribution in Social Policy," in *The Philosophy of Welfare*, 213.

³⁸² Titmuss, "The Role of Redistribution in Social Policy," 211.

well aware of the role of class politics in achieving this goal, admitting that “[f]rom some perspectives these major changes in policy could be regarded as ideological pleas to the middle- and upper-income classes to share in the benefits (as well as the costs) of public welfare.”³⁸³ The National Health Services, for example, had succeeded in providing “non-discriminatory, non-judgmental” care in large part because the middle classes joined in 1948 and remained thereafter in the system instead of contracting out.³⁸⁴

What Titmuss had less successfully appreciated, by his own later admission, was the fact that social insurance *qua* insurance could neither fully escape class politics nor as a result achieve the solidaristic aims he had associated with it. Instead, insurance as tracking both earned reward and basic need persisted precisely because it avoided the full application of either principle. This feature was particularly important in winning over the most important constituency for broad welfare policies, the middle class. As historian Peter Baldwin has explained,

To the extent that social policy was to be more extensive, inclusive, or generous...it no longer reapportioned resources *de haut en bas*, but increasingly within one large middle group. Growing more powerful, the interests of this new constituency began to determine the flow of redistribution.... In nations where statutory intervention was accepted as normal and desirable...it did not take long for the European welfare state to formulate, as among its main tasks, the concern of the middle class for itself.³⁸⁵

Universal welfare policies therefore succeeded not by overriding actuarial reasoning but by appealing to it, and in particular by tapping into the desire of the bourgeoisie to secure itself without sacrificing too much to the poor, with whom it felt it had little in common. In this respect there was indeed a serious tension between the aims of social solidarity, which sought to move beyond the limited

³⁸³ Ibid.

³⁸⁴ Ibid., 216.

³⁸⁵ Baldwin, *The Politics of Social Solidarity*, 30-1.

identity or equality of restricted risk pools, and the distributive politics underlying the post-war welfare state. As a result, social insurance soon came under attack from both ends of the political spectrum, as commentators alleged that the rubric of risk tended to conceal or limit the true aims of social welfare.

C. The limits of universal social insurance

I have been suggesting that if there was a change in welfare policies between the pre- and post-war period, it was in the scope of their application rather than the theoretical principle at work. Yet we have also just seen that the goal of encompassing social solidarity pointed beyond what the insurance idea, even in its most expansive form, was authorized or able to do. The result was a system that depended on actuarialism in practice—specifically, on the self-aware solidarity of the middle classes—while abjuring it in principle. Indeed, the political success of Beveridge’s proposals hinged largely on their ability to conceal the extent to which they deviated from actuarialism, preserving the aura of traditional insurance while obscuring the amount of vertical redistribution (that is, redistribution from the wealthy to the poor) involved. By the same token, post-war attempts that were more explicit about their redistributive intentions failed due to resistance from the property-owning middle classes.³⁸⁶ Universal social insurance, in short, was palatable primarily insofar as it appealed to class-based self-interest and hid the ways in which it betrayed that appeal.³⁸⁷

³⁸⁶ In the decades after the war, risk-pooling in France was confined to groups of social and actuarial peers, and in Germany social insurance remained the province of workers already included in social welfare schemes. Only later on, when demographic and economic changes weakened the petit bourgeoisie in these countries did one-time objectors to universal welfare demand and achieve broader risk-pooling that would accommodate them as well. *Ibid.*, 206. In Finland, to take just one more example, health insurance was blocked during the inter-war period by the Agrarian party, which refused to pay for urban benefits, and was extended to cover the entire population only in 1962. Cutler and Johnson, “The Birth and Growth of the Social Insurance State,” 102.

³⁸⁷ In the United States, for example, old-age security insurance is paid for by a highly visible payroll tax, which creates resistance to benefits increases and vertical redistribution. By contrast, in Sweden and other European countries, pensions are financed partly by employers and partly out of general tax revenues, both of which allow policies’

The claim that there is something inherently deceptive or unduly restrictive about social insurance consequently became part of a prominent critique of the practice in the second half of the 20th century.³⁸⁸ Libertarians argued that it was really forced income redistribution under the guise of a contract, while socialists took issue with the gap between its solidaristic rhetoric and its ability to lift the poor. Although the flexible or prudential character of risk-pooling had originally been one of the political selling points of social insurance, it came during the 1950s and 60s to appear to many as a liability, concealing or limiting the moral aims undergirding welfare policy.

A pioneer of this critique was again Burns, who in a 1953 article argued that the confusion between social and traditional insurance was creating serious problems in political discourse and public policy. The inclusion within social insurance schemes of non-risks such as the existence of children; the increasing detachment of benefits from contributions; and the reliance on tax and employer financing had made actuarialism a distant reality. As a result, “in most countries with extensive and highly-developed social security systems, social insurance has become an institution to which the word ‘insurance’ can be applied only at the risk of a serious distortion of the language.”³⁸⁹ The danger, according to Burns, is more than linguistic: Because the process of crafting social policy is an inherently political one, involving conflicting interests and objectives, private insurance concepts “not only provide no solution, but may have the dangerous consequence of partially concealing the fact that these issues and conflicts exist.”³⁹⁰

distributive effects to be concealed to the insured to a greater extent. See John B. Williamson and Fred C. Pampel, *Old-Age Security in Comparative Perspective* (New York and Oxford: Oxford University Press, 1993), 104.

³⁸⁸ See Eveline Burns, “Private and Social Insurance and the Problem of Social Security,” *Canadian Welfare* 28, no. 7 (1953): 5-10 and Richard Titmuss, “Models of Redistribution in Social Security and Private Insurance,” in *Commitment to Welfare* (London: George Allen & Unwin Ltd., 1968), 184.

³⁸⁹ Burns, “Private and Social Insurance,” 6.

³⁹⁰ *Ibid.*, 9.

A few years later, Friedrich Hayek cited Burns's work in making similar point about the obfuscating character of the insurance model. While private group insurance cannot in principle effect a deliberate transfer income from one risk group to another, he argued, social security systems had come over time to make vertical redistribution their explicit goal. "No system of monopolistic compulsory insurance has resisted this transformation into something quite different, an instrument for the compulsory redistribution of income.... It was mainly through decisions that seemed to most people to concern minor technical issues, where the essential directions were often deliberately obscured by an assiduous and skillful propaganda, that the transformation was effected."³⁹¹ One consequence, according to Hayek, is an increasing gulf between the actual workings of welfare schemes and their public perceptions, a "serious problem" for democracy given that "the immense social security apparatus has been a chief factor in the transformation of our economy," yet is "also the least understood."³⁹² The new power of central governments to divide "income according to some preconceived notion of justice" had been granted unsuspectingly by democratic publics thanks to the "stroke of promotional genius" of calling tax- and debt-financed redistribution "insurance."³⁹³

Even Titmuss, who had initially supported universalism to remove the stigma of poor relief, came by 1965 to lament the use of insurance principles in social welfare. Like Hayek, Titmuss allowed that in principle the goal of equity in private group insurance entails a deliberate rejection of redistribution: Participants in such schemes see themselves as having entered into an individual contract, in which their contributions earn no more and no less than their own actuarially determined benefits. In practice, however, Titmuss conceded that all group insurance, both private

³⁹¹ F. A. Hayek, *The Constitution of Liberty*, in *The Collected Works of F. A. Hayek, Vol. XVII*, Ronald Hamowy, ed. (Chicago: University of Chicago Press, 2011), 409-410.

³⁹² *Ibid.*, 410.

³⁹³ *Ibid.*, 427 and 409, citing Lewis Meriam and Karl Schlotterbeck, *The Cost and Financing of Social Security* (Washington, DC: Brookings Institution, 1950), 8.

and social, is redistributive to the extent that its underlying risk ratings are imprecise, a reality that had become increasingly evident over time. The problem then is not the occurrence of vertical redistribution, as Hayek claimed, but rather the fact that without proper awareness and planning, both of which the insurance rubric seems to undermine, such redistribution may end up being regressive rather than serving the poor.³⁹⁴ Titmuss concluded that the concept of insurance in social welfare should be abandoned altogether in favor of universal contributions, flat-rate benefits, general tax financing, and selective benefits to the needy as a matter of right.³⁹⁵

It has thus long been understood to both the partisans and the opponents of the egalitarian welfare state that the rubric of risk and insurance may conceal as much as it illuminates about the workings of these programs and the normative aims behind them. As Burns once again presciently put it, “the institution of social insurance is a social invention which was brought into being to perform a specific function in a specific economic and social environment.”³⁹⁶ Its resemblance to private insurance allowed for a shift away from individual economic responsibility in certain cases, serving in effect as a bridge between a more traditional liberal ethos and a newer social one. Yet while the private analogy “enabled social insurance to perform this task, its very success carried within it the seeds of its own destruction, or at least very fundamental modification.”³⁹⁷

What these observers may have failed to understand is the difficulty of extricating the insurance idea from the politics of the liberal welfare state. The Webbs and Beveridge were right that the public’s receptivity to insurance as a mean of sorts between distributive extremes would allow governments to achieve more expansive aims by blurring the lines of actuarial fairness. Yet

³⁹⁴ Titmuss, “Models of Redistribution in Social Security and Private Insurance,” 175-184.

³⁹⁵ Ibid., 184. See also his “Universal and Selective Social Services” (1967) and “The Role of Redistribution in Social Policy” (1964) in *Commitment to Welfare*, 113-23 and 196-97, respectively.

³⁹⁶ Eveline M. Burns, “Social Insurance in Evolution,” 199.

³⁹⁷ Eveline M. Burns, “Social Security in Evolution: Toward What?,” *Social Service Review* 39, no. 2 (1965): 129.

the shared perception of vulnerability and risk—often, but not necessarily, as aligned with economic class—remained a critical source of public support for welfare policies, and this foundation became if anything more important as democracy spread over the course of the 20th century.³⁹⁸ The politics of welfare thus remained closely tied to class-based risk pooling even as the normative theory tried to move beyond it.

I now wish to switch gears somewhat, to consider contemporaneous developments in mathematical probability and economic thought. We will see that at around the same time that proponents of universal welfare sought to transcend the limits of empirical risk pooling in the name of a classless egalitarianism, a newly subjective understanding of probability undermined the case for risk pooling from the other direction. Adherents to the subjective view of probability focused on economic rationality and saw risk-taking as an omnipresent but wholly private affair. Their accounts therefore abandoned the political understanding of insurance—and in particular, its dual-faceted distributive appeal—just as welfare egalitarians did the same.

II. Subjective probability and the personalization of chance

A. Keynes's transitional account

Probability, as we have seen, has always been subject to two distinct interpretations, an epistemic and an aleatory. The difference between them became the focus of sustained philosophical consideration in conjunction with the emergence of frequentism, which in its most rigorous form insists that mathematical probabilities exist empirically and bear no relation to the

³⁹⁸ This is the compellingly argued thesis of Baldwin's *The Politics of Social Solidarity*. See also Titmuss, "The Role of Redistribution in Social Policy," 196. For a more recent iteration, with a different empirical argument, see Iversen, *Capitalism, Democracy, and Welfare*.

individual's state of mind.³⁹⁹ The alternative view, that probability is a measurement of partial belief, had been central to the understanding of the term prior to the emergence of the probability calculus in the 17th century.⁴⁰⁰ It continued to play a role in mathematical theory through Laplace and his successors, as long as theorists assumed that subjective likelihoods (that is, equiprobability assignments under the principle of insufficient reason) were the necessary starting point for determining objective ones (via inverse probability and the principle of succession). After the split that occurred thanks to frequentism, however, the epistemic view, while famously articulated by Augustus De Morgan in 1847, was not seriously considered again until the early decades of the 20th century.

One of the first thinkers to reconsider the subjective interpretation of probability was John Maynard Keynes. Along with Beveridge, Keynes is widely considered a founding father of the modern welfare state.⁴⁰¹ He is also, if less famously, a notable figure in the history of probability theory. Probability was the subject of his Cambridge fellowship dissertation and of his first major book, *A Treatise of Probability*, published in 1921. Although his direct influence on the field was ultimately limited, Keynes reflects a broader shift that took place around this time between the

³⁹⁹ Keynes rightly finds the distinction present even in Hume: "Probability is of two kinds, either when the object is really in itself uncertain, and to be determined by chance; or when, though the object be already certain, yet 'tis uncertain to our judgment, which finds a number of proofs on each side of the question." Hume, *A Treatise of Human Nature*, Book ii, part iii, section ix, cited in Keynes, *Treatise*, 322. Within mathematical probability theory proper, the honor of first making this distinction goes to Laplace, although he conflated the two views in practice. As Keynes rightly notes, Cournot was the first to consistently differentiate between them.

⁴⁰⁰ An overview of the non-mathematical understandings of probability at this time can be found in Casson, *Liberating Judgment*.

⁴⁰¹ For instance, historian Hugh Heclo avers that two episodes of social learning were essential to the emergence of modern centralized welfare: first, an understanding that unemployment is a function of macroeconomic causes rather than a lack of individual responsibility; and second, the discovery of techniques of aggregate demand management that allowed for the continuity in the provision of social services even during periods of economic recession. The latter in particular paved the way for a full-scale welfare state, and is attributable directly to Keynes's economic work. Heclo, *Modern Social Politics*, ch. 3. Keynes, incidentally, was also the Treasury official responsible for negotiating pension spending with Beveridge. *Ibid.*, 256.

frequentist interpretation that had been dominant in England since roughly the final third of the 19th century and the subjective one that would largely supplant it over the course of the 20th.⁴⁰²

Keynes was not a subjectivist in the sense that would later prevail, and which we will consider in greater detail in the next two subsections. Yet he did understand probability as concerned first and foremost with the conditions of individual belief. It is “‘subjective chance,’ concerned with knowledge and ignorance,” that is “‘fundamental,’” he explained, and “‘so-called ‘objective chance,’ however important it may turn out to be from the practical or scientific point of view, is really a special kind of ‘subjective chance’ and a derivative type of the latter.”⁴⁰³ Pursuant to this understanding, Keynes took aim at frequentism and proposed a new account of probability as a logical relation between propositions. His account is logical, he explained, “because it is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals...”⁴⁰⁴ It differs from most subsequent subjective theories in maintaining that only certain classes of likelihoods are quantifiable, but shares with them a focus on systematizing the process of inference by which all probability values are obtained.

The contrast between Keynes and Edgeworth, who traveled in the same circles and shared many interests, is particularly illuminating of the changing philosophical tide that interests us here. As we saw, Edgeworth had turned to statistics in the years after publishing *Mathematical Psychics* with the aim of furthering his project of a scientific utilitarianism. As historian Steven Stigler points out, his early statistical essays contain several notable innovations, including a version of an argument for

⁴⁰² Contemporary Bayesians may not be quick to recognize the significance of Keynes in this respect, since his own view of statistical inference was ultimately quite skeptical. See, e.g., Christian P. Robert, “Reading Keynes’ *Treatise on Probability*,” *International Statistical Review* 79, no. 1 (2011): 1-15. Yet, as we will see below, Bayesian decision theory would not have been thinkable without the advent of modern subjective probability, and in this respect, if not in his actual statistical thinking, Keynes represents a turning point.

⁴⁰³ Keynes, *Treatise*, 327.

⁴⁰⁴ *Ibid.*, 3.

the central limit theorem and an early appearance of the Student's t distribution.⁴⁰⁵ The articles that immediately followed these, on the best estimate of a mean and its use in testing the significance of statistical data, served, in Stigler's words, as "basic references for the theory and application of statistical techniques to social and economic data" until the end of the century.⁴⁰⁶

Keynes would later charge that Edgeworth's mature focus on the applications of statistics came at the expense of his earlier concern with the philosophical foundations of the discipline. In his obituary for his senior colleague, Keynes recounted that when he pressed Edgeworth to "give an opinion as to how far the modern theory of statistics and correlation can stand in the frequency theory falls as a logical doctrine," the latter

would always reply to the effect that the collapse of the frequency theory would affect the universality of application of statistical theory, but that large masses of statistical data did, nevertheless in his opinion, satisfy the conditions required for the validity of Statistical Theory, whatever these might be. I expect that this is true. It is a reasonable attitude for one who is mainly interested in statistics to take up. But it implied in Edgeworth an unwillingness to revise or take up again the more speculative studies of his youth.⁴⁰⁷

Thus whereas Edgeworth generally accepted the "rule that probability implies reference to a series," Keynes charged that the frequentist is fatally unable to justify his choice of any particular reference class.⁴⁰⁸ Whereas Edgeworth maintained that it was possible to defend the credibility of probability's initial assumptions, including the *a priori* uniformity of unknown distributions, Keynes argued that such justifications are either empirically unfounded, failing on the frequentist's own criteria, or

⁴⁰⁵ Stephen M. Stigler, "Francis Ysidro Edgeworth, Statistician," *Journal of the Royal Statistical Society* 141, no. 3 (1978): 294.

⁴⁰⁶ *Ibid.*, 297.

⁴⁰⁷ J. M. Keynes, "Francis Ysidro Edgeworth, 1845–1926," *The Economic Journal* 36, no. 141 (1926): 147–48. See also John Aldrich, "Keynes among the Statisticians," *History of Political Economy* 40, no. 2 (2008): 287–88.

⁴⁰⁸ Edgeworth, "The Philosophy of Chance (II)," 260, 275; Keynes, *Treatise*, 114.

circular.⁴⁰⁹ And whereas Edgeworth continued to believe that probabilistic expectation could guide individual conduct, Keynes expressed grave concern about its relevance, arguing that it fails among other things to capture the actual psychology of risk-taking and the relevant particulars of individual cases.⁴¹⁰

Keynes was forceful on many of these points, particularly the applicability of expectation to personal choice, as even Edgeworth seemed at one point to admit.⁴¹¹ But by casting doubt on the foundations of statistics without offering a compelling positive theory or any new statistical techniques of his own, he consigned himself to relative obscurity within the discipline.⁴¹² This fact alone offers a window onto the inherently practical bent of mathematical probability, which Keynes seemed eager to avoid.⁴¹³ He ventured into the traditional moral territory of the field only fleetingly,

⁴⁰⁹ Edgeworth, “The Philosophy of Chance (II),” 263-6, 272-3; Keynes, *Treatise*, 118-20. Perhaps most significantly, Edgeworth concluded that Keynes’s skepticism on this point did not penetrate to the law of errors, on which the former pinned most of his hopes for the contribution of statistics to the social sciences.

⁴¹⁰ Keynes offers three arguments concerning the relevance of probability and in particular classical expectation to individual conduct: first, that it assumes, along with utilitarian ethical theory, that degrees of goodness are numerically measurable and additive; second, that it ignores the “weights” of arguments or the evidence on which each probability is founded; and finally that it ignores the element of “risk,” or in other words how great or small is the amount at stake in the uncertain event. See Keynes, *Treatise*, 355-369. The last two points are explicitly reminiscent of D’Alembert’s critique, which we encountered earlier. In addition, as has already been stated, Keynes holds that not all probabilities are quantifiable at all, in which case if “the question of right action is under all circumstances a determinate problem, it must be in virtue of an intuitive judgment directed to the situation as a whole, and not in virtue of an arithmetical deduction derived from a series of separate judgments directed to the individual alternatives each treated in isolation.” *Ibid.*, 357.

⁴¹¹ Edgeworth, “The Philosophy of Chance (II),” 278-9: “In private life occasions occur requiring that we should act for the nonce so to speak, or at least without anticipating a succession of similar actions which would afford a practical certainty of advantage in the long run. In public life the same choice may occur.... Utilitarian philosophy has been too silent about this difficulty.”

⁴¹² On the reception of Keynes’s work and his legacy (or lack thereof) in the field of statistics, see Aldrich, “Keynes among the Statisticians.”

⁴¹³ Edgeworth, in his review, affirms as much when, analogizing the question of a probability’s “weight” to the theory of errors, he explains that it is preferable “with Laplace to seek, not that combination of the given observations which is most probably...*right*, but the one which, minimising the detriment incident to the use of fallible observations, maximizes the Expectation of *useful results*.” Edgeworth, “The Philosophy of Chance (II),” 275 (italics mine). Edgeworth also, on similar grounds, condoned the assumption that the normal law of errors is fulfilled in cases where precise numerical measurement is not possible. “Always presuming that the normal law is in a sense fulfilled, it appears possible to determine an exact quantitative correlation...in respect of such unquantified attributes, as eyecolour or good temper. The practice of the classical writers on Probabilities who did not hesitate to make ‘moral’...advantage a subject of calculation, seems to be countenanced by this modern art of measurement.” Edgeworth, “On the Use of the Theory of Probabilities,” 184.

when he suggested—partly in a footnote—that gambling is objectionable because it entails an individual loss in the long run and is likely to increase inequality in the short run. He immediately retreated from expanding on these claims, however, since it “would lead too far from what is relevant to the study of Probability.”⁴¹⁴

In this respect Keynes placed himself outside of the tradition that we have been considering until now. He did not have any particular regard for insurance as a successful application of probability theory, and even mildly chastised the industry for encouraging the misguided “presumption in favour of the numerical valuation” of all likelihoods.⁴¹⁵ At the same time, his insistence on the fundamentally epistemic character of probability anticipated the view that would come to dominate economic thinking in the 20th century, and his relative lack of interest in its moral, legal, and political implications reflects this new orientation. In this regard, then, Keynes may be seen as a harbinger of the broader turn away from the collectivist view of probability that had resonated with utilitarian political economy and helped support earlier welfare practice. As we will now see, this turn had significant consequences for thinking about social welfare in the second half of the 20th century.

B. The rise and triumph of the epistemic view

The modern theory of subjective probability was inaugurated in earnest in 1926, with Frank Plumpton Ramsey’s essay “Truth and Probability,” which argued that “the laws of probability are

⁴¹⁴ Keynes, *Treatise*, 367.

⁴¹⁵ *Ibid.*, 22. On the political front, Keynes reportedly regarded Beveridgean social insurance—unfunded, non-actuarial schemes that make no attempt to adjust premiums to individual risks or to align contributions with benefits—as “an empty fiction,” which merely “concealed a convenient form of taxation.” Beveridge, by contrast, “believed that the form of insurance had an inherent political virtue.” Karel Williams and John Williams, “Social Insurance and the Allied Services: The Political Utopia of 1942” in *A Beveridge Reader*, ed. Karel Williams and John Williams (London: Allen & Unwin, 1987), 47.

laws of consistency, an extension to partial beliefs of formal logic, the logic of consistency.”⁴¹⁶

Ramsey was certainly not the first to try to formulate a coherent account of inductive inference with reference to the probability calculus, and wrote this article in the wake of important contributions to the same effort by Peirce as well as Keynes. Yet while Peirce was a resolute objectivist, arguing that the probability of an argument’s truth is a long-run empirical frequency, and while Keynes devised his own idiosyncratic theory of probabilities as logical relations, Ramsey’s account of probabilistic inference was entirely personalistic, resting on initial degrees of belief that are set independently of evidence or any other external standard. “...[T]o ask what initial degrees of belief are justified...seems to me a meaningless question; and even if it had a meaning I do not see how it could be answered.”⁴¹⁷ Instead, the subjective theory would focus on the nature of rational inference from initial beliefs or “priors” that are simply taken as givens.⁴¹⁸

Ramsey based this account of rational inference on a series of hypothetical bets. To have “any definite degree of belief implies a...willingness to bet on a given proposition at the same odds for any stake,” as well as “a consistency between the odds acceptable on different propositions as shall prevent a book being made against you.”⁴¹⁹ Beliefs are thus consistent if one has a clear set of preferences that do not change depending on the order in which they are presented, or in other words if they follow the axioms of probability. If this is not the case, Ramsey shows, then the individual “could have a book made against him by a cunning better and would then stand to lose in

⁴¹⁶ Frank Plumpton Ramsey, “Truth and Probability,” in *Studies in Subjective Probability*, ed. Henry E. Kyburg, Jr. and Howard E. Smokler (New York: Robert E. Krieger Co., 1980), 41.

⁴¹⁷ *Ibid.*, 48.

⁴¹⁸ See H. Kyburg, “Subjective Probability: Criticisms, Reflections, and Problems,” *Journal of Philosophical Logic* 7, no. 1 (1978): 158.

⁴¹⁹ Ramsey, “Truth and Probability,” 42.

any event.”⁴²⁰ This reliance on betting as the standard for rational belief “will not seem unreasonable when it is seen that all our lives we are in a sense betting.... The options God gives us are always conditional on our guessing whether a certain proposition is true.”⁴²¹

In contrast to both the frequentist and logical accounts, then, the subjective interpretation of probability tends to hold that all likelihoods are quantifiable: Their numerical value is simply a measure of the individual’s personal belief in an event or statement given the evidence available to him.⁴²² Subjective theories also differ from earlier accounts in maintaining internal coherence as their only demand on an individual’s beliefs, meaning that in order to act rationally in situations of uncertainty, it is sufficient to avoid having a distribution of degrees of belief that will entail a likely or certain loss of utility.⁴²³ With Bruno de Finetti’s introduction in 1931 of the concept of exchangeability or symmetry, which characterizes events the estimated probability of which does not depend on the order of their appearance, it became possible to connect these subjective probability values with the procedures of Bayesian statistical inference, in which one starts with an initial probability distribution and modifies it in light of experience. De Finetti showed that for a series of exchangeable events, a reasonable individual will arrive after sufficient observation at a probability value close to the event’s observed frequency regardless of the opinions he held at the outset.

Because of this universal compatibility with Bayes’ theorem, the subjective interpretation of probability quickly gave rise to an entire field of analysis known as Bayesian decision theory, which

⁴²⁰ Ibid., 41.

⁴²¹ Ibid., 42.

⁴²² This is not uniformly true of all subjective accounts: Some followed Keynes in allowing for degrees of belief that are incomparable to others in their order of magnitude. For a summary see Kyburg and Smokler, “Introduction,” in *Studies in Subjective Probability*, 12-13.

⁴²³ Ibid., 8. On weak accounts of coherence, the bettor is incoherent if he bets in such a way that he will inevitably lose money on the wager; on strong accounts, he is incoherent if he bets in such a way that there is a chance he will lose or come out even. Ibid., 14.

applied inverse statistical methods to generate an account of rational decision-making based on an individual's subjective priors.⁴²⁴ The Bayesian decision maker learns from empirical observation, incorporating an objective element into the theory. Yet the foundations of his beliefs—that is, the initial probabilities from which learning begins—are still entirely personal to him, and as a result the equilibrium or long-run probabilities of decision-makers will not necessarily converge.⁴²⁵ In other words, an account of probabilistic belief that begins with wholly subjective likelihoods is far less likely to provide a basis for the kind of solidaristic risk pooling that we saw emerging from the frequentist account.

By mid-century, many mathematicians and economists in this tradition had come to embrace the roles of risk and insurance in distributive policy. As Kenneth Arrow put it in a seminal article on the economics of healthcare, “a good part of the preference for redistribution expressed in government taxation and expenditure policies and private charity can be reinterpreted as desire for insurance.”⁴²⁶ In noting that government support tends to go “to those who are disadvantaged in life by events the incidence of which is popularly regarded as unpredictable,” Arrow was echoing the probabilistic argument for social insurance as we have been tracing it since the late-18th century.⁴²⁷ Yet because of the radically subjective foundations of their approach, the “risk” entailed in mid-20th century economic arguments for welfare was not the same “risk” that had inspired the earliest

⁴²⁴ See Edi Karmi and David Schmeidler, “On the Uniqueness of Subjective Probabilities,” *Economic Theory* 3, no. 2 (Apr. 1993): 268.

⁴²⁵ This point is given illustration in Richard Breen and Cecilia García-Peñalosa, “Bayesian Learning and Gender Segregation,” *Journal of Labor Economics* 20, no. 4 (2002): 899-922 and Thomas Piketty, “Social Mobility and Redistributive Politics,” *The Quarterly Journal of Economics* 110, no. 3 (1995): 551-584.

⁴²⁶ Kenneth J. Arrow, “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review* 53, no. 5 (1963): 947.

⁴²⁷ *Ibid.*, 947-48.

welfare programs or, by imperfect extension, the post-war welfare state. Let us now consider in more detail how these accounts differed.

C. The moral character of subjective probability

Like all accounts of probability, the subjective view lends itself to a strongly normative interpretation. Specifically, it posits coherence as a necessary condition of rationality, and rationality as the ability to achieve one's desired ends or maximize personal utility.⁴²⁸ Unlike the classical and frequentist views, however, this particular normative approach discarded the explicit concern of probability theory with contractual equity and focused instead on "rational betting," or the "avoidance of certainty of losing to a clever opponent."⁴²⁹ Thus where earlier accounts took great pains to distinguish the counsels of mathematically enlightened reason from gambling, the subjective view embraced the wager-like character of all decisions.

This shift was in part a consequence of the behaviorist assumptions of subjective probability theory: In order to infer an individual's personal probabilities and utilities, one must start with a preference ranking of acts, all of which involve uncertainty in the same way and which may therefore be considered as fundamentally alike. Although accounts of subjective probability were not the only theories of rational choice to infer psychological states or preferences from wagers—John von Neumann and Oskar Morgenstern famously did this in their *Theory of Games and Economic Behavior* as well—they were the first to do so, and as such paved the way for the normalization of betting as a reflection of acceptable and even reasonable predilections.⁴³⁰

⁴²⁸ On the failure of subjective accounts to offer a descriptive theory of decision-making, see Kyburg, "Subjective Probability," 165-66.

⁴²⁹ R. Sherman Lehman, "On Confirmation and Rational Betting," *Journal of Symbolic Logic* 20, no. 3 (1955): 251.

⁴³⁰ This trend received explicit confirmation in a very important article by Milton Friedman and Leonard Savage, which used the von Neumann-Morgenstern utility analysis to describe how risk "preferences" of various sorts can explain (i.e.,

This change in the normative character of probability theory had, it seems to me, two consequences that are directly relevant to the theory of social insurance. The first was an unlimited proliferation of risks, understood as any quantified likelihood or prediction, and with it of the conceivable range of insurance-like products to manage them. As Arrow put it, “the variety of possible risks in the world is really staggering. The relevant commodities include, in effect, bets on all possible occurrences in the world which impinge upon utilities.”⁴³¹ Second, while the applicability of insurance as a “commodity” spread, mutual insurance as an ethical practice lost its unique significance and appeal. As we have seen, philosophers of probability since Daniel Bernoulli had assigned a pride of place to insurance as one of the most important applications of the calculus, diametrically opposed to gambling in its personally and collectively beneficial effects.⁴³² On the subjective theory, however, probabilities lost much of their interpersonal relevance and contractual fairness or justice became a secondary concern, subordinate to the demands of personal utility maximization on the basis of risk preferences taken as givens.

This is not to say that advocates of the subjective or Bayesian approach were unaware of the need for a theory of “multisubjective” decisions, as de Finetti termed them, which would allow for

rationalize) both gambles and insurance purchases. Milton Friedman and L. J. Savage, “The Utility Analysis of Choices Involving Risk,” *Journal of Political Economy* 56, no. 4 (1948): 279-304. On this analysis, distributive claims become predilections or aversions to risk, as captured by a utility curve: An unskilled worker, for example, “may jump at an actuarially fair gamble that offers a small chance of lifting him out of the class of unskilled workers and into the ‘middle’ or ‘upper’ class, even though it is far more likely than [another] gamble to make him one of the least prosperous unskilled workers. Men will and do take great risks to distinguish themselves, even when they know what the risks are. May not the concave segment of the utility curve...translate the economic counterpart of this phenomenon appropriately?” *Ibid.*, 299.

⁴³¹ Arrow, “Uncertainty and the Welfare Economics of Medical Care,” 945.

⁴³² Assuming a concave utility curve, Bernoulli showed, even mathematically fair wagers entail an expected loss to both players. Insurance, by contrast, can be framed as a decision *not* to wager, to secure for oneself a certain sum rather than exposing the value of one’s goods and thereby facing a loss the utility of which might exceed their mathematical value (depending on one’s initial wealth). D. Bernoulli, “New Theory on the Measurement of Risk,” 23-36. Bernoulli neglects to explain why the insurer’s action is not akin to a wager, or why it is that the underlying economic activity (involving the marine transport of goods for sale) is not itself a form of gambling given the possibility of loss. Nevertheless, his notion of moral expectation and the assumption of money’s diminishing marginal utility remained mainstays of probabilistic attempts to distinguish the two practices.

the drawing of valid social conclusions irrespective of the parties' different premises.⁴³³ As Arrow put it in a 1958 review article, the results of statistics "are supposed to be interpersonally valid. It follows that even if one accepts the subjective or personal probability viewpoint as a guide to an individual's action, there is need for some method of combining the utility functions and subjective probabilities of different individuals."⁴³⁴ Yet as Arrow readily acknowledged, decision theorists could not reach a consensus regarding the best criterion for social choice under uncertainty, as each available alternative involved serious flaws.⁴³⁵ It is also striking that the two social decision rules that most resemble mutual insurance—the minimax and maximin theorems, both of which dictate minimizing one's potential losses—eschew the language of insurance completely, to say nothing of mutuality or sympathy between the parties.⁴³⁶ It is fair to say that the subjective theory of risk demoted the effort to align probability's epistemic and aleatory sides, an effort that had defined the practical character of the calculus since its inception.

⁴³³ Bruno de Finetti, "Recent Suggestions for the Reconciliation of Theories of Probability," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley: University of California Press, 1951), 220.

⁴³⁴ Kenneth J. Arrow, "Utilities, Attitudes, Choices: A Review Note," *Econometrica* 26, no. 1 (1956): 8.

⁴³⁵ The question of social choice under conditions of uncertainty remains a subject of debate in welfare economics. See, e.g., Marc Fleurbaey, "Assessing Risky Social Situations," *Journal of Political Economy* 118, no. 4 (2010): 649-680 and Matthew D. Adler and Chris William Sanchirico, "Inequality and Uncertainty: Theory and Legal Applications," *University of Pennsylvania Law Review* 155, no. 2 (2006): 279-377. Cf. Amartya Sen, *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1970), 140: "The theory of decision-taking under uncertainty does not yield very definite conclusions on problems of this kind."

⁴³⁶ For instance, L. J. Savage had reinterpreted the minimax theorem, originally the equilibrium solution to a two-person game, to apply to conditions of uncertainty. This rule counsels the individual to choose the act that leads to the smallest possible loss, with loss understood as the difference between the income she could obtain if she knew the actual state of the world (or "nature") and the income she does obtain by choosing any particular outcome without knowing that state. In a group context, where members assign different likelihoods to the possible states of nature, the rule dictates that "an act be adopted such that the largest loss faced by any member of the group will be as small as possible." Yet while this principle allows for a group decision that is independent of the parties' probability assessments, Savage explains, it also ensures that the parties' opinions have no real influence on one another or even on the decision itself. In addition, while the accumulation of further evidence will reduce parties' expected loss under the rule, it bears no necessary relation to the certainty with which each holds his own views and typically does not bring them into greater agreement with one another. Thus Savage's minimax solution suggests that a collective decision based on personal likelihoods will be extremely limited in its aims, as participants seek not to persuade one another of the truth but rather to simply guard their individual interests against the "risk" that someone else's judgment will triumph. Leonard J. Savage, *The Foundations of Statistics*, 2nd revised ed. (New York: Dover, 1972), 173-75. See also Arrow, "Utilities, Attitudes, Choices," 11-12.

As a result, subjective accounts of probability also undermined the kind of probabilistically informed fellow feeling that had inspired earlier arguments for mutual and social insurance. Even if risk pooling could be justified as a reflection of personal risk aversion, and social insurance as a correction of market failures, the perception of equal or shared vulnerability that had long been central to the case for those practices was far less accessible under the subjective account and its decision-theoretic successors.⁴³⁷ I now wish to show that the most influential work of political philosophy to emerge from this period, Rawls's *Theory of Justice*, can be understood precisely as an attempt to rescue social insurance, and with it the liberal welfare state, by doing without probabilities what neither welfare theorists nor economists could do with them.

III. The egalitarian welfare state without probability

A. Probability versus justice

We have seen that by the mid-20th century the virtues of probabilistic social insurance had been called into question on two fronts. On one side, the insurance rubric and its de facto actuarialism had proven themselves to be in tension with the egalitarian aims of post-war welfare. On the other side, subjective probability theory had normalized gambling and undercut the case for solidaristic risk pooling. If the frequentist interpretation of probability had encouraged a strong identification with the collective—the idea that, in Peirce's words, “our interests...must not stop at our own fate, but must embrace the whole community”—the subjective view shrunk the scope of

⁴³⁷ Indeed, there is an argument from decision theory that mutual insurance will not arise at all under conditions of uncertainty, because in order to perceive mutual burden-sharing as fair the parties must assume that their odds of encountering a given outcome are equal. Only by relying on conditions external to the theory, such as familiarity, trust, or social control, can the subjectivist convincingly account for such an assumption. See Göran Skogh, “Risk-Sharing Institutions for Unpredictable Losses,” *Journal of Institutional and Theoretical Economics* 155, no. 3 (1999): 510-12.

each individual's interest down to his own case alone.⁴³⁸ It may not be a coincidence, then, that important strands of social and political theory gave up on risk at right about the time that risk itself stopped being social or political.

Whatever the relationship between these developments, Rawls's recasting of the social contract as a problem of decision-making under conditions of uncertainty can be understood as a reaction to both. First, his rejection of all probabilities in the original position allowed him to distance the choice procedure from a gamble and to align it instead with a kind of insurance, wherein the individual sets out to guarantee his basic security or minimize his likely exposure to misfortune. Second, such insurance without probabilities could in turn allow for precisely the kind of broad-based solidarity that the post-war welfare state had tried but failed to achieve.

Rawls was not the first 20th-century thinker to devise an account of collective choice based on a hypothetical scenario of ignorance.⁴³⁹ Yet he was the first, as far as I am aware, to adopt the maximin criterion as the best choice for participants.⁴⁴⁰ This strategy—borrowed from game theory, where it is the equilibrium solution to the two-person, zero-sum game—dictates focusing on one's worst possible payoff and choosing the option that maximizes it.⁴⁴¹ Unlike a game situation,

⁴³⁸ Like Venn, Peirce maintained that because “the idea of probability essentially belongs to a kind of inference which is repeated indefinitely,” it follows “that there can be no sense in reasoning in an isolated case, at all.” Peirce, “On the Doctrine of Chances, with Later Reflections,” 161-2.

⁴³⁹ For precedents, see William Vickrey, “Measuring Marginal Utility by Reactions to Risk,” *Econometrica* 13, no. 4 (October 1945): 319-33 and “Utility, Strategy, and Social Decision Rules,” *The Quarterly Journal of Economics* 74, no. 4 (November 1960): 507-35, and John Harsanyi, “Cardinal Utility in Welfare Economics and in the Theory of Risk-taking,” *Journal of Political Economy* 61, no. 5 (October 1953): 434-35.

⁴⁴⁰ For the purposes of this discussion, I will not distinguish between the original maximin rule and its lexicographic variation, which Rawls adopted in response to criticism that the maximin does not always satisfy strong Pareto optimality. Both assume that the individual decision-maker will focus first and foremost on the worst-case scenario. See Amartya Sen, *Collective Choice and Social Welfare*, 138, and “On Weights and Measures: Informational Constraints in Welfare Economics,” *Econometrica* 45, no. 7 (October 1977): 1546.

⁴⁴¹ The maximin's most extreme alternative, the maximax of gain, calls for the player to always choose the gamble whose best prize is highest, regardless of what he stands to lose in doing so. The minimax of regret focuses on minimizing the maximum possible opportunity cost of an incorrect decision, on the assumption that the decision-maker wants more than anything to avoid regretting his choices. There are also mixed strategies, which let a random device assign probabilities and calculate expected utilities on that basis. The decision-maker's attitude toward risk is also reflected in

however, in which the individual faces a calculating opponent whose goal is to outsmart him, in a decision context one's "opponent" is nature, which may be arbitrary but is not unfailingly wicked. The choice of this strategy in a decision context has therefore been described as "a manifestation of pure cowardice," focusing so intently on the worst-case scenario that it neglects all other possible outcomes.⁴⁴² Indeed, Rawls admits that the principles of justice representing the maximin strategy are "those a person would choose for the design of a society in which his enemy is to assign him his place."⁴⁴³

These two principles are, first, that each person has an equal right to the maximum basic liberty compatible with the liberty of others; and second, that social and economic inequalities are attached to open positions and offices, and arranged so that they redound to the advantage of all. The latter, known as the difference principle, is what ensures that inequalities of fortune are not allowed to become sources of dependency or self-contempt.⁴⁴⁴ It represents "an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be."⁴⁴⁵ This image of a shared resource pool, the products of which are to be divided fairly in accordance with the presumed *ex ante* equality of participants, bears a striking resemblance to the original probabilistic model of contractual equity. The main difference is that here the absence of all mathematical likelihoods enables a much more complete assumption

the utility he assigns to payoffs whose likelihood is unknown. See William Fellner, *Probability and Profit: A Study of Economic Behavior along Bayesian Lines* (Homewood, IL: Richard D. Irwin, 1965), 142.

⁴⁴² William J. Baumol, *Economic Theory and Operations Analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1961), 370.

⁴⁴³ John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 152.

⁴⁴⁴ See also John Rawls, "Reply to Alexander and Musgrave," *The Quarterly Journal of Economics* 88, 4 (1974): 641. "Particularly important are the effects [of the principles of justice] on the social bases of self-respect; for when self-respect is lacking, we feel our ends not worth pursuing, and nothing has much value." Compare Burns, "Priorities for Public Welfare," 39, noting that the major advantage of social insurance is its complete avoidance of means tests, for "it makes little difference to one's sense of self-respect whether one has to submit to this procedure to obtain the whole of one's monthly income or to obtain the missing 8 or 10 percent."

⁴⁴⁵ Rawls, *A Theory of Justice*, 101.

of equality, and with it greater receptivity to comprehensive mutual insurance of a kind, than any agreement based on calculated risks could achieve.

Rawls explains the choice of the maximin principle as a direct consequence of the absence of probabilities in the original position. The veil of ignorance, as he sets it up, ensures that the parties have “no basis for determining the probable nature of their society, or their place in it.” Because such “knowledge of likelihoods is impossible, or at best extremely insecure,” the participants have “strong reasons for being wary of probability calculations if any other course is open to them.”⁴⁴⁶

This assumption of probabilistic ignorance is necessitated above all by Rawls’s premise that “injustice exists because agreements are made too late,” after “people already know their social positions and relative strength in bargaining.”⁴⁴⁷ A contract based on empirical probabilities will therefore merely reflect existing morally irrelevant advantages, as participants segregate themselves by category and jockey for resources on that basis. Subjective likelihoods are unavailable as well, since probability estimates should have an “objective basis...in knowledge of particular facts” if they are to serve as public reasons.⁴⁴⁸ As Rawls was surely well aware, subjective probabilities had been shown in other accounts of social choice to have little if any interpersonal authority, even if they could be the basis for certain (flawed) collective decisions.

B. Rawls’s defense of social insurance

Despite this rejection of probabilities, however, Rawls was apparently determined to preserve the insurance-like character of the parties’ choice. The other two features of the decision procedure that justify the maximin strategy, he explains, are that the alternatives to the guaranteed

⁴⁴⁶ Ibid., 154-55.

⁴⁴⁷ John Rawls, “Some Reasons for the Maximin Criterion,” *The American Economic Review* 64, no. 2 (1974): 141.

⁴⁴⁸ Rawls, *A Theory of Justice*, 172.

minimum are unacceptable—that is, another conception of justice could result in “serious infractions of liberty for the sake of greater social benefits”—and that the person choosing “has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend.”⁴⁴⁹ Rawls denies that individuals in the original position are inherently risk-averse, but rather takes it as his task to show that choosing “as if one had such an aversion is rational given the unique features of that situation.”⁴⁵⁰ Parties behaving in this way will choose not only to “protect their basic rights but [also to] insure themselves against the worst eventualities,” including the loss of their own freedom for someone else’s advantage.⁴⁵¹

This last point in particular distinguishes justice as fairness from utilitarianism, which in adopting “for society as a whole the principle of rational choice for one man” entails what is effectively a gamble.⁴⁵² As Rawls explained in a subsequent article, while it may be the case that a utilitarian principle will sometimes secure basic individual liberties and interests, “there is no reason why it will do so in general,” and “it would be pointless to run the risk of encountering circumstances when it does not.”⁴⁵³ This association between utilitarianism and gambling also helps to explain why Rawls rejected another social decision rule available under conditions of uncertainty: the equiprobability assumption proposed by John Harsanyi.⁴⁵⁴ Harsanyi reinterpreted the classic principle of indifference (according to which one assumes equal probabilities in the absence of knowledge to the contrary) as an expression of ethical impersonality, allowing the individual to

⁴⁴⁹ Ibid., 154, 156.

⁴⁵⁰ Ibid., 172.

⁴⁵¹ Ibid., 176.

⁴⁵² Ibid., 26-7. Indeed, Rawls cites directly to F. Y. Edgeworth’s *Mathematical Psychics* as an example of a work deriving the principle of utility from contract theory. Ibid., 29.

⁴⁵³ Rawls, “Some Reasons for the Maximin Criterion,” 143.

⁴⁵⁴ Harsanyi, “Cardinal Utility in Welfare Economics,” 434-35.

imagine himself as interchangeable with everyone else in his society. This is of course quite similar to Rawls's depiction of the veil of ignorance, except that under the equiprobability assumption, Harsanyi shows, participants will choose to maximize the overall sum of utilities, since each one's expectation is an equal portion of the total pool.⁴⁵⁵ The rule is thus insensitive to both individual liberties and to *ex-post* individual welfare levels, and as such represents precisely the kind of wager that Rawls believes his participants would reasonably avoid.⁴⁵⁶

It is quite clear, then, that on Rawls's view every available account of probabilities poses a problem for distributive justice. In a later article, he explained that because "there is at present so much disagreement about the meaning and interpretation of probability," arguments from the parties' ignorance may be best used to support the "strains of commitment" case for maximin. This is the view that "no one is permitted to agree to a principle if they have reason to doubt that they will be able to honor the consequences of its consistent application." When choosing between two possible accounts of justice, one of which would permit or require social positions that the individual could not accept, while the other results in "arrangements that everyone can honor in all circumstances," the decision-maker must agree to the latter. The two principles, which "always secure acceptable conditions for all," are therefore preferable to any criterion that may not.⁴⁵⁷ In short, whatever account of probabilities one chooses, reasoning based on mathematical likelihoods could place one in a position that is too difficult to bear—a risk that no reasonable person should take.

⁴⁵⁵ This point is also made by Sen, *Collective Choice and Social Welfare*, 141-2.

⁴⁵⁶ Ibid. 142-4 and Peter A. Diamond, "Cardinal Welfare, Individualistic Ethics, and the Interpersonal Comparison of Utility," *Journal of Political Economy* 75, no. 5 (1967): 765-6.

⁴⁵⁷ Rawls, "Reply to Alexander and Musgrave," 652.

I therefore propose that we understand Rawls's defense of the maximin rule first of all as an effort to rescue the logic of social insurance from the countervailing pulls of contemporary probabilistic and economic thought. Even in conditions of uncertainty, Rawls tells us, it is possible to make a decision that is not in essence a wager on whether a certain empirical proposition is true. Rather, once empirical or probabilistic reasoning is ruled out, Rawls shows that the personally and collectively beneficial outcome derives from each person's choice to guarantee for himself an acceptable minimum even under the worst-case scenario—in other words, to insure on the assumption of complete aleatory equality or interchangeability.

This defense of non-actuarial social insurance can in turn also be read as an astute commentary on the politics of post-war social welfare. If the impetus for the egalitarian welfare state was, in Richard Titmuss's words, "connected with the demand for one society; for non-discriminatory services for all without distinction of class, income or race; for services and relations which would deepen and enlarge self-respect; for services which would manifestly encourage social integration," Rawls suggests that the only way to achieve this in practice is by so deeply insinuating the insurance idea into the fabric of social relations that the resulting arrangement ceases to outwardly resemble insurance. When it comes to not only the basic guarantees of liberty and security but also the foundations of self-respect, Rawls explains, no one should choose to gamble.⁴⁵⁸ In this, he echoes the liberal moralists of the probabilistic tradition, who argued for insurance as the personally responsible choice and as a commitment to the overall social order. Unlike previous probabilistic accounts, however, justice as fairness maintains the liberal character of a fair contract while doing away with the particularistic distributive entitlements that had allegedly limited the scope

⁴⁵⁸ On the centrality of self-respect to post-war welfare theory, see also J. Donald Moon, "The Moral Basis of the Democratic Welfare State," in *Democracy and the Welfare State*, ed. Amy Gutmann (Princeton, NJ: Princeton University Press, 1988), 27-52.

of earlier welfare. The same extreme uncertainty that mandates caution also guarantees inclusiveness, Rawls shows, because it unsettles traditional distributive politics completely.

C. Insurance and distributive theory after Rawls

I have suggested that understanding both the limitations of post-war social insurance and the way in which 20th-century economic thought undermined the practice sheds light on Rawls's attempt to formulate a theory of distributive justice that, while closely resembling a polity-wide insurance agreement, jettisoned probabilities even more completely and explicitly than universal welfare programs had done. If my interpretation is right, then Rawls's theory also implies a critique of any continued reliance on traditional insurance principles as the foundation of welfare policies. Contemporary thinkers such as Robert Goodin have argued that the political virtue of understanding welfare transfers as insurance is that "redistribution, of a sort, is thus justified without any appeal to old-style and increasingly unfashionable values of equality or altruism." Thus the "solidarity" of the shared risk pool," along with the efficiencies of public over private insurance provision, "may be quite enough to motivate support for something rather like the welfare state into the indefinite future."⁴⁵⁹

This assessment, however, does not acknowledge the central problem we have been considering: that of defining the "shared risk pool" in a way that is simultaneously prudential and redistributive. The reintroduction of means tests into many Beveridge-type systems over the second half of the 20th century points up the limits of an approach that aspires to universality while relying

⁴⁵⁹ Robert E. Goodin, "The End of the Welfare State?," 216. Jacob Hacker defines a social welfare program as one that "exists to protect individuals from widespread risks inherent in the market economy," and notes that because occupational benefits also "encompass elements of risk sharing and cross-subsidization," they too should be understood as part of the welfare state. Hacker, *The Divided Welfare State*, 32-3.

on the perception of contractual fairness for political support.⁴⁶⁰ Indeed, if the politics of welfare retrenchment of the 1980s are any indication, it is precisely universal insurance programs on the post-war model that often prove most susceptible to reform pressures arising from both fiscal and moral concerns.⁴⁶¹ These programs, which tend to provide only minimal benefits, are limited in their impact on the poor, insufficient for the upwardly-mobile middle classes (who often turn to private insurance and fringe benefits to supplement their coverage), and, where relatively more generous, vulnerable to the critique that they discourage individual responsibility.⁴⁶² It seems that Rawls has a powerful case that only by doing away with risk and actuarialism is anything like universal, egalitarian welfare sound in principle and attainable in practice.⁴⁶³

Yet while Rawls's theory seems to preserve many of the virtues of social insurance—its responsiveness to both desert and need, its liberal yet egalitarian character—without its apparent

⁴⁶⁰ On the introduction of means-tested public assistance within many insurance-based schemes see, e.g., E. Philip Davis, *Pension Funds, Retirement-Income Security, and Capital Markets: An International Perspective* (Oxford: Clarendon Press, 1995), 40.

⁴⁶¹ Paul Pierson, *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment* (Cambridge: Cambridge University Press, 1994), 57, 102-3.

⁴⁶² On the tendency of Beveridge-style programs to give rise to a kind of dualism rather than egalitarianism as the middle classes opt out or supplement their coverage, see Esping-Andersen, *The Three Worlds of Welfare Capitalism*, 25.

⁴⁶³ Elizabeth Anderson's impassioned plea for earnings-related social insurance relies heavily on class-based support, and in particular a presumed middle-class preference for such systems over minimum-benefits programs. Because "low, flat benefits are too minimal to appeal to the middle classes, who can self-insure at the low levels promised," and "[g]raded benefits add security against loss of middle-class standing," "Bismarckian" systems are more appealing to middle-class voters and socially more beneficial overall than their "Beveridgean" counterparts. See Elizabeth Anderson, "Common Property," *Boston Review*, July 25, 2016. Anderson's argument, while commendable on many levels, seems to overlook two important points. First, it has historically been the case that non-democratic governments are more likely than democratic ones to adopt earnings-related systems, while democratic countries are quicker to adopt minimal and means-tested programs. See Cutler and Johnson, "The Birth and Growth of the Social Insurance State," 112. Indeed, as Beveridge's own views attest, there is an affinity between a system of minimal state provision, supplemented by communal and familial aid, and the so-called bourgeois or middle-class virtues of hard work and self-reliance. Given this, it appears that Anderson's middle-class voters are either unaware of their true interest in graded social insurance or warier than she admits of how much such programs will have to deviate from actuarial fairness in order to benefit the poor. Second, and closely related to the foregoing, Anderson does not adequately explain how earnings-related insurance would help the neediest given that the "size of their claims is based on impersonal rules graded to their prior contributions into the system." Absent means tests, graded systems have to deviate considerably from traditional insurance principles in order to help the poor. Where the political support for such generosity would come from remains an open question.

limitations, it leaves us with the question of what exactly we give up when we exclude probabilities from the equation. By precluding the empirically grounded sense of identity and the stronger but more circumscribed solidarity of the earlier frequentist approach—and, to an even greater extent, its friendly society precursors—does his account not undercut one of the welfare state’s most important sources of psychological and political support? Perhaps such spontaneous, self-identified risk groups are the closest a liberal welfare regime can come to the kind of uncalculating relationship that points beyond justice. At any rate, they come closer to doing so than anything Rawls seems to envision. If social insurance can be both individually choiceworthy and universal in scope only if it is not probabilistic, we would do well to ask whether choosing the first two features over the third does not undermine the purpose and the viability of the welfare project itself.

The same question arises with respect to another account of distributive justice that emerged in Rawls’s wake, and which also invoked insurance as a rational choice under conditions of uncertainty. Luck egalitarians take their bearings from Ronald Dworkin’s distinction between “brute” and “option” luck, or between bad fortune that is beyond the individual’s control and inequality that results from deliberate choices or gambles. The rationale is that, according to the ruling principle of equality, “it is unjust when some people lead their lives with less wealth available to them, or in otherwise less favorable circumstances, than others, not through some choice or gamble of their own but through brute bad luck.”⁴⁶⁴ Dworkin therefore envisions the ideal distribution as emerging out of a hypothetical insurance market, in which each participant starts with equal resources and, based on the assumption that he is equally likely as everyone else to encounter a given misfortune, is offered the chance to purchase reasonably priced insurance against that eventuality. The resulting arrangement, Dworkin concludes, will compensate for bad luck that is

⁴⁶⁴ Ronald Dworkin, *Sovereign Virtue* (Cambridge, MA and London: Harvard University Press, 2002), 347.

beyond the individual's control while allowing for differences in outcome that emerge from personal preferences and choice.

Dworkin makes a number of important assumptions about this hypothetical insurance market, which takes slightly different forms depending on the risk in question but preserves the same overall character for each. The first is that participants will regard themselves as equally likely to encounter the given risk or eventuality. Although he admits, in the context of unemployment, that this is a “nearly impossible assumption,” it is apparently necessary for the second critical feature of the market, namely that it will offer coverage “at community rates,” or at the same price for the same coverage for everyone.⁴⁶⁵ Because “risks of most catastrophes are now regarded by the actual insurance market as randomly distributed,” and because the assumption of equal likelihoods precludes further classifications, insurers will not be able to discriminate against those more susceptible, genetically or otherwise, to harm.⁴⁶⁶ Finally, Dworkin maintains that the resulting benefits will preclude the kind of discontent that arises when some people feel the undeserving are “unfairly capitalizing on [their] efforts.”⁴⁶⁷ Those who receive compensation “would have paid a premium reflecting the cost of that option to others,” while those who experience poor option luck thanks to their unwillingness to insure will have to acknowledge their own responsibility for their fate.⁴⁶⁸

In essence, then, Dworkin is proposing an account of social insurance, a scheme for mutual protection based on what “the average member of the community” would “supposedly” have

⁴⁶⁵ Ibid., 332.

⁴⁶⁶ Ibid., 78.

⁴⁶⁷ Ibid., 332.

⁴⁶⁸ Ibid.

chosen “if circumstances had been more equal than they are.”⁴⁶⁹ Unlike Rawls, however, he does not appear to be particularly concerned with the difficulties of probabilistic reasoning and their potentially confounding role in his vision. Instead, his hypothetical insurance markets incorporate both an epistemic interpretation of probability, in participants’ assumption of their equal likelihoods, and an aleatory one, in the form of the presumably random distribution of risks. Without the latter condition, it would not be possible to imagine a market for insurance at all, since insurers would not be able to predict the likelihood of harm and adjust their finances accordingly. Without the former, the market would be highly segmented and some people might be priced out.

This combination of arguments accords with certain intuitions about fairness, and seems to avoid some of the pitfalls of both Harsanyi’s utilitarian application of the principle of indifference and Rawls’s difference principle.⁴⁷⁰ Yet it is also vulnerable to objection from each of its two probabilistic sides. First, on the epistemic side, Dworkin’s design of the decision procedure and his endorsement of the principle of indifference presume that it makes sense to assign probability values to singular events. It is not obvious why we should subscribe to such a view, either as a matter of logic or in light of its political implications. Even if we do accept it, however, we are left with the question of why decision-makers—especially those who already know their “hopes, fears, tastes, and values,” and in some cases their talents and ambitions as well—would assume their equal susceptibility to everyone else, or be reasonable to do so.⁴⁷¹

⁴⁶⁹ Ibid., 78.

⁴⁷⁰ Dworkin’s critique of both alternative principles can be found in *ibid.*, 328-31.

⁴⁷¹ Ibid., 332. Hopes and fears in particular bear quite a strong relationship to probability estimates, as Hume recognized: “The passions of fear and hope may arise when the chances are equal on both sides, and no superiority can be discovered in the one above the other. Nay, in this situation the passions are rather the strongest, as the mind has then the least foundation to rest upon, and is toss’d with the greatest uncertainty. Throw in a superior degree of probability to the side of grief, you immediately see that passion diffuse itself over the composition, and tincture it into fear. Encrease the probability, and by that means the grief, the fear prevails still more and more...” David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2000), 283.

Second, on the aleatory side, it is unclear how an insurance market could survive as a market on the admittedly inaccurate assumption that the risks in question are randomly distributed across the entire population. Dworkin acknowledges the problem of adverse selection, which would induce insurers to raise rates in response to the over-subscription of bad risks, and says in the context of healthcare that the only way to prevent this is to nationalize insurance.⁴⁷² Yet such insurance, in order to provide the same coverage to everyone despite their empirical differences, would either have to offer only low benefits across the board or to subsidize those with higher risk profiles.⁴⁷³ In the first case, the scheme may offer less coverage than the average person would choose, and as a result fail to reflect the “true measure of the social resources devoted to the life of one person” as determined by “how important, in fact, that resource is to others.”⁴⁷⁴ In the second case, it is susceptible to the objection of unfairness, since it would have to treat those whom it insists are empirical equals in obviously unequal ways.

Ultimately, then, it would appear that the epistemic and aleatory assumptions of Dworkin’s scheme are in tension, and not simply congruent as he implies. Either probabilistic equality represents a moral constraint, a kind of ethical impartiality, or it is a statement about the world. If it is the former, then it need not be justified on empirical grounds, but it does not provide much guidance for individual decisions in anything resembling a realistic insurance market. If the latter, it permits the kind of insurance that Dworkin has in mind, but only insofar as risks are in fact

⁴⁷² Dworkin, *Sovereign Virtue*, 452.

⁴⁷³ Indeed Dworkin acknowledges that insurance against risks such as underemployment and health will be provided at relatively low levels under his scheme. See *ibid.*, 94-99, 314-16. He makes this argument as a function of what individuals with concave utility curves would choose, however, and not as a result of what insurance providers would have to offer given the limitations of the market as he conceives it. My point here is that the level of coverage that insurance companies or the state could provide on Dworkin’s account may be even lower than what the average person would choose, given the difficulties of working under such non-actuarial assumptions.

⁴⁷⁴ *Ibid.*, 70.

randomly distributed across the entire population, and we know that this is not the case for every risk.

This conclusion reflects the enduring difficulty we have been considering throughout this dissertation, namely that of theoretically reconciling probability's two sides. Dworkin's insistence on their compatibility seems to reflect his desire to vindicate liberal equality as the sovereign political good, presiding over a harmonious vision in which "liberty, equality, and community are not three distinct and often conflicting political virtues," but "complementary" and impossible to understand in isolation from one another.⁴⁷⁵ Social insurance has historically reflected this kind of harmonizing impulse, and luck egalitarianism is apparently no exception. The very idea of a distribution that is at once "ambition sensitive," reflecting individual judgments and choices in the face of uncertainty, and "endowment insensitive," compensating for statistically defined harms, is yet another attempt to square the probabilistic circle. But it is not clear that any form of insurance can achieve such harmony in principle. Rather, our sustained consideration of probability theory implies that it is only possible in practice, and only approximately at that. Without this awareness, we may neglect both the psychological sources on which social insurance depends and the careful balancing act that it unavoidably entails.

IV. The fate of social insurance in the 20th century and beyond

In the first part of this chapter, I argued that Beveridge-style social insurance failed to accomplish its goal of adequate universal coverage, in large part because of the probabilistic psychology on which it rested. While aspiring to universal self-respect and social cohesion, it depended critically on the participation of the middle classes, and in particular on members'

⁴⁷⁵ Ibid., 237.

perceptions of their own similarity and of the risks that they as a group were liable to face. As a result its generosity toward the worst off was limited, and eventually its programs had to be supplemented or supplanted by targeted and means-tested programs for the poor—precisely the type of differentiated approach that advocates of universal welfare had sought to avoid.⁴⁷⁶

This failure of social insurance to realize the inclusive aims of its early advocates became, by the middle of the 20th century, part of a powerful critique of the idea among welfare theorists. I have interpreted Rawls's *Theory of Justice* as part of this line of critique, arguing that social insurance *as* insurance—that is, as a contract reflective of both the individual's knowledge of his own situation and as a tool for redistributing burdens across a group—cannot be universal in its reach as long as it remains tied to probability estimates of any sort. Yet I have also suggested that Rawls undervalues the psychological underpinnings of the welfare state, and in particular the sense of identity and sympathy on which the political demand for social insurance rests. To be sure, such identity is shifting and contingent; its generosity is bounded by perceptions of similarity; it can be divisive and even dangerous if unmoored by a concern for the common good. Yet it reflects a kind of mutual concern and reciprocity that points toward friendship on a national political scale. By excluding such empirical affinities from their accounts of liberal equality, both Rawls and his luck egalitarian successors preclude the brand of solidaristic reasoning on which, I have argued, the welfare state was built.⁴⁷⁷

The alternative justification for social insurance that emerged in the last century was based on its ability to satisfy preexisting consumer preferences and to correct failures of the private market

⁴⁷⁶ Earnings-related (or what is now popularly known as “Bismarckian”) social insurance is in this regard more forthright, since in grading contributions to benefits it explicitly differentiates between classes of wage-earners rather than using the contractual appearance of insurance primarily as a guise for vertical redistribution. Yet such graded coverage is not egalitarian in the way that Beveridge, Titmuss, or Rawls envisioned, and also raises the objection that it fails to help those who need it most.

⁴⁷⁷ They also rule out the basis of much contemporary electoral support for welfare, which recent scholarship has shown to be highly correlated with shared perceptions of risk. See Rehm, Hacker, and Schlesinger, “Insecure Alliances.”

to do the same. This account also differed significantly in its underlying psychology from the one that had inspired earlier welfare programs and even, if only partially, the universal welfare state. For one, it took both individual attitudes toward risk and personal probability estimates as givens, inserting a radical subjectivity into the foundation of its analysis. While contemporary Bayesians do allow for the adjustment of personal probabilities in light of experience, the arbitrariness of an individual's starting point and the limited extent of each person's experiential learning make it unlikely that anyone's estimates will ultimately align with anyone else's, particularly if they begin with very different views.⁴⁷⁸ As a result, these arguments undermine the possibility of a political or public-minded argument for mutual insurance, based on the understanding that one's own fate is tied to another's and on a feeling of responsibility to one's extended "own." The subjective account gave up on the effort to defend, let alone instill, a broadened view of individual concern.

The twentieth century consequently witnessed the development of two opposing, even polarizing intellectual tendencies with regard to social insurance and the welfare state more generally. On one side, advocates of a robust social equality sought a moral case for welfare that did away with empirical risk assessments and the particularistic claims to which they give rise. On the other side, advocates of a contractual, market-oriented approach set out to create a system that would reflect individual preferences and subjective judgments. These two dispositions align for the most part with the political left and right, respectively, and their difference of principle is in some degree to be expected. The left, with its focus on universal needs and vulnerabilities, emphasizes equality and commonality, while the right seeks to preserve a kind of hierarchy determined by personal choice and ambition. Yet the heightened divide between these two views in contemporary politics may also

⁴⁷⁸ See Breen and García-Peñalosa, "Bayesian Learning and Gender Segregation" and Piketty, "Social Mobility and Redistributive Politics."

reflect the difficulty, in light of 20th-century developments, of articulating an account of social insurance that can accommodate both.

A more compelling theory of social insurance would recognize that there are conflicting virtues at work in any such distributive scheme. Probabilistic reasoning cannot resolve this conflict, but it can tutor or elevate each side. It can help the partisans of distinction recognize their shared vulnerability, not only on the level of family and community—as today’s conservatives insist—but also on a larger and more encompassing scale. At the same time, it can remind advocates of abstract equality of the psychological basis of solidarity, which begins with a perception of one’s own empirical situation and extends outward from there. We have seen that probability has always concerned itself with both aspects of uncertainty: the statistical view that sees the individual as just one event in a larger series, difficult if not impossible to comprehend in isolation; and the epistemic view that allows for personal judgment based on the particularities of one’s own case. Social insurance is the practical reflection of the effort to harmonize these two perspectives.

The challenge, then, for those who seek to expand the scope of redistribution beyond limited risk pools is not to deny the relevance of empirical probabilities, but rather to persuade would-be distributive equals of their shared vulnerability and interest in mutual protection. At the same time, it will be important to recognize that this type of solidarity is inherently constrained, both by existing knowledge about risks and by the claims of distinction or ambition that encourage people to imagine themselves as above average. Social insurance will therefore have to be supplemented by other arguments—from collective benefit, from personal virtue, from compassion or altruism—as one element of a political theory of distributive justice. Nevertheless, when properly understood it provides an appealing framework for mediating distributive conflicts: a stable if somewhat shifting foundation on which many basic needs of a population can be met and, at the same time, its members’ own judgments and aspirations given reign.

Conclusion

I. The longstanding appeal of insurance

The preceding chapters have traced a conceptual history of social insurance, from the origins of risk through the advanced welfare state. I argued that probability theory played a central role in the development of welfare thinking and practice, implying an account of distributive justice that has proven remarkably influential in political life. Probabilistic justice, as I have been calling it, had the potential to replace divisive claims of desert and need with an account of equality that in some way encompassed both. Insurance understood as a political principle therefore promised to replace longstanding distributive disputes and their accompanying passions with impartial determinations of equality based on the rigorous counsels of mathematics.

In making this case, I advanced three principal claims about probability theory and its relationship to welfare thinking. The first is that mathematical probability is inherently and inescapably normative in its character. We saw that the very project of quantifying likelihoods grew out of a moral and legal question, the need to apportion fair shares in an interrupted aleatory contract. Each subsequent account of probability has in turn both reflected and furthered the practical aims of its exponents. This should not be surprising, given that the discipline is at its core an attempt to quantify good judgment and equity, neither of which can be understood in isolation from the ends to which they point. The assumptions and techniques of mathematical probability are therefore intimately bound up with its exponents' views about the character and aims of practical reason, including its relationship to justice broadly understood.

The second claim in my argument, which follows directly from the first, is that probability theory is fundamentally concerned with the problem of reconciling individual choice with some account of the common good. It was not long after the founding of the discipline that probabilists began to realize the potential disconnect between contractual equity, as defined by their calculations,

and individual prudence or common sense. Much of the subsequent development of the theory can be understood as a series of attempts to resolve this problem. Each account had a different character and resulted in different proposals. Yet they held in common the promise of harmonizing instrumental reason with fairness or justice, which roughly correspond to the two sides of probability itself.

Finally, I have argued that the answers to this problem that emerged within and out of probability theory from roughly the end of the 18th through the 20th century were deeply influential, in ways that have not been sufficiently remarked, to the development of the modern welfare state. Statistical insurance was the first practice in which philosophers of probability sought, and in their view found, the means to reconcile individual benefit with the common good. The application of insurance principles on a social scale therefore promised to extend such harmony well beyond isolated associations to politics as a whole. Insurance would reflect the free choices of individuals while simultaneously securing social order. It would give each citizen his due while promoting the aggregate benefit. And it would distribute resources on the basis of both personal responsibility and shared need, accommodating the two principles without clearly favoring either one.

As it turns out, the practical success of insurance in this regard far exceeded its theoretical capacity. That is, it was not by theoretically resolving the conflict between the individual and collective good that social insurance established itself in political life and remained salient over the course of roughly two centuries. Rather, it did this by subsuming competing principles within the technical, impartial language of risk. Indeed, the fact that social insurance has remained the dominant paradigm for welfare provision despite historically embodying very different assumptions and aims speaks to its flexibility and resulting political appeal. Insurance manages to reconcile competing principles by concealing, at least in part, the irresolvable conflict between them.

Probability theory set the stage for this coup, in that it advanced an account of distributive justice based on what appear to be impartial, unassailable claims of equality. In fact, determinations of probabilistic equality are inherently partial, for two reasons. First, they require the decision maker to classify events, whether *a priori* or in light of observation, by choosing some characteristics as relevant and disregarding others. Second, and even more fundamentally, they are relative to the account of probability in light of which they are made. Both types of partiality are largely hidden to the untrained eye, whether behind statistical regularities or in technical devices such as the principle of indifference. They affect the distributive character of any social insurance program while remaining removed from the realm of open political conflict.

This interpretation of insurance as a somewhat deceptive distributive regime is intended not simply as a critique, but also as a reflection on the important mediating function that insurance has served in modern politics. Social insurance is the direct heir to the liberal project: a contract, frequently hypothetical, made among those who regard themselves as equally vulnerable to some harm, which sets out to guarantee their individual and collective security. Like liberalism itself, it has spawned countless criticisms and undergone various modifications. Its character has changed as the understanding of risk evolved from an individual entitlement attached to a social obligation, to the property of a collective, to a wholly personal prediction. Yet throughout these evolutions, it has served the same overarching purpose, with varying degrees of success: namely, to unite individuals with diverse natures and interests under a single principle of reward or recompense.

II. Explaining the welfare state

It has not been my intention here to offer an exhaustive explanation for the emergence and evolution of the modern welfare state, a question beyond the scope of such a work and in any case better left to historians and empirical political scientists. Nevertheless, my argument does address a

puzzle in the study of welfare policy, namely, why the welfare state continues to be associated among both scholars and the general public with insurance, despite the tenuous relationship between many welfare policies and traditional insurance practices.⁴⁷⁹ It also suggests a corrective to some recent scholarship in the field.

Since roughly the middle of the 20th century, it has been common in economics, political science, and legal studies to explain social insurance policies as promoting economic efficiency, specifically by correcting market failures and reducing the social cost of compensating for accidents.⁴⁸⁰ Even when it departs from actuarial principles, on this view, compulsory or publicly orchestrated insurance performs the needed function of covering risks that are inadequately addressed by private markets and therefore likely to impose externalities or impede overall economic growth.⁴⁸¹ Such efficiency-based explanations certainly help to explain why policymakers might opt for mandatory coverage instead of leaving certain risks to voluntary arrangements. From the point of view of the social planner, it is desirable to minimize the overall cost of accidents and to improve aggregate economic performance even if doing so requires deviating from strict actuarial entitlements. These accounts may also explain why democratic and especially middle-class voters—who are often critical for the enactment of welfare legislation—would support welfare programs despite their vertically redistributive character. Where private markets fail to provide adequate or affordable protection, as has long been the case for risks such as unemployment, sickness, and

⁴⁷⁹ Some contemporary scholars do not even notice the inconsistency of referring to wholly non-actuarial programs as “social insurance.” See, e.g., Cutler and Johnson, “The Birth and Growth of the Social Insurance State,” 89.

⁴⁸⁰ For just a small sample, see Kenneth J. Arrow, “Uncertainty and the Welfare Economics of Medical Care”; Guido Calabresi, *The Cost of Accidents* (London and New Haven: Yale University Press, 1970), esp. ch. 4; Nicholas Barr, “Social Insurance as an Efficiency Device,” *Journal of Public Policy*, Vol. 9, No. 1 (1989): 59-82.

⁴⁸¹ See Barr, “Social Insurance as an Efficiency Device,” 66.

retirement, citizens may be prudent to support public programs that protect them even at a cost that exceeds their actuarial expectations.⁴⁸²

Explanations for welfare that focus solely on individual economic advantage, however, do not account for the claims of mutual obligation or solidarity that have inspired so many political and philosophical accounts of the welfare state. In this regard, the class-mobilization thesis now prominent in comparative welfare-state studies comes closer to understanding social insurance as the complex political phenomenon that it is.⁴⁸³ More specifically, a recent strand of thinking that understands the preference for welfare as a demand for insurance echoes the frequentist view I described in chapter three and offers a compelling account of the link between electoral politics and social-policy outcomes.⁴⁸⁴ *Ex ante*, welfare policies provide voters with a desired sense of protection; *ex post*, they redistribute wealth from economic winners to losers.⁴⁸⁵ Political scientists working in this tradition have set out to show how preferences rooted in voters' economic positions are aggregated into social policies that in turn fill both insurance and redistributive functions.

⁴⁸² Ibid.

⁴⁸³ The seminal text of this school is Esping-Andersen's *The Three Worlds of Welfare Capitalism*. One may disagree, however, with Esping-Andersen's claim that "social citizenship constitutes the core idea of the welfare state." Esping-Andersen, *The Three Worlds of Welfare Capitalism*, 21. The history of social insurance gives us no particular reason to prioritize the post-war, solidaristic or de-commodifying variety of welfare to any of those that came before it. In fact, if we regard social insurance as the unchanging core of the modern welfare state, the social-citizenship model may be as much a departure from the original welfare idea as its continuation. Torben Iversen's remark that the "welfare state is simultaneously an arena for distributive struggles *and* a source of comparative advantage" is more in keeping with the conceptual analysis of social insurance I have offered here. Iversen, *Capitalism, Democracy and Welfare*, 13.

⁴⁸⁴ See Iversen, *Capitalism, Democracy and Welfare*; Baldwin, *The Politics of Social Solidarity*. It might be open to some debate whether these are two distinct schools of thought or rather the second is a novel interpretation of the first. See Rehm, Hacker, and Schlesinger, "Insecure Alliances," characterizing the latter as "revisionist." Given that Iversen and others accept Esping-Andersen's focus on the relationship between welfare and electoral politics, I have characterized them here as cousins. See Torben Iversen and David Soskice, "Politics for Markets," *Journal of European Social Policy* 25, no. 1 (2015): 76-93. One important point of distinction, however, is the latter school's willingness to question the power-resources view that welfare capitalism is "primarily about decommodification and exploitation of the rich." Iversen, *Capitalism, Democracy and Welfare*, 7.

⁴⁸⁵ This explanation, as we have seen, can be traced back at least to Pigou, *Wealth and Welfare*, 366-69.

This explanation indeed helps to account for how many welfare policies were enacted in advanced democracies, especially in the period just prior to and since the Second World War. It also indicates that efficiency- and solidarity-based arguments for welfare can be, and often have been, closely aligned. Policies that protect against the dangers of income disruption or destitution can reflect both a rational response to market-generated insecurity and a form of solidarity, binding together groups of citizens who perceive themselves as similarly vulnerable. Although the demand for redistribution reflects self-interest, it also necessitates joining with others and, on the basis of presumed affinities, creating broader electoral coalitions. Probabilistic solidarity based on personal-empirical risk assessment therefore combines a self-regarding rationale with one that is, by necessity, oriented toward the group. Accounts of social choice that rule out such probabilistic reasoning also deny the kind of electoral coalitions that have in fact been responsible for welfare provision in many countries since the turn of the 20th century.

At the same time, my argument suggests that scholars of the welfare state ought to pay greater attention to the nature of social insurance as a regime—a mixed regime, one could say, at least at its best—and not merely as a response to the self-seeking claims of particular groups. The history of probability reveals that insurance can accommodate a variety of different answers regarding the definition of equality to which distributive claims should correspond.⁴⁸⁶ If indeed the political genius of social insurance lies in its ability to accommodate more than one such answer,

⁴⁸⁶ There is a relatively recent social science literature on this question as it applies to private insurance, and specifically the dilemmas of risk classification as a potentially discriminatory practice. See Kenneth S. Abraham, “Efficiency and Fairness in Insurance Risk Classification,” *Virginia Law Review* 71, no. 3 (1985): 403-451; Ine Van Hoyweghen, Klasien Horstman and Rita Schepers, “Genetics is not the issue’: insurers on genetics and life insurance,” *New Genetics and Society* 24, no. 1 (2005): 79-98; Yves Thiery and Caroline Van Schoubroeck, “Fairness and Equality in Insurance Classification,” *Geneva Papers on Risk and Insurance* 31(2006): 190-211; and Jyri Liukko, “Genetic discrimination, insurance, and solidarity: an analysis of the argumentation for fair risk classification,” *New Genetics and Society* 29, no. 4 (2010): 457-475. Although these works concern private rather than social insurance practice, their focus on the definition of probabilistic equality necessarily highlights the continuities between the two. As one author puts it, “the boundaries between private and social insurance have always been unclear...and the recent discussion has questioned the distinction even further.” Liukko, “Genetic discrimination,” 458.

both at once and over time, political scientists and others would do well to analyze the mixed distributive consequences of social insurance programs as part of their design and staying power. Doing so would make clear not only the sources of political support for welfare programs, but also the distinct and at times conflicting moral principles that animate them.

III. A word for contemporary partisans

I have been characterizing social insurance as essentially a moderating force, capable of accommodating and reflecting competing distributive claims. This is a rather different view of it from the caricatures offered in contemporary political debates. Ultimately, I hope my argument will enlighten partisans on both sides by showing that social insurance speaks to the concerns of each, but not quite in the way that true believers might hope.

When today's conservatives advocate private savings and economic mobility as the cure to social ills, they are in essence endorsing a distributive regime that rewards a certain kind of virtue.⁴⁸⁷ When progressives talk about shared responsibility and social equality, they are suggesting among other things that distribution should track basic needs, both physical and psychological. We have seen that social insurance satisfies both principles, albeit only partially and provisionally. On one hand, it appears to reward individual foresight and economic success; on the other, it rests on a feeling of identification rooted in shared vulnerability. This quality, which has made social insurance so appealing, has also guaranteed that it would remain forever subject to competing interpretations and struggles. Advocates of distinction and desert will prefer restrictive risk pools, limiting their responsibility to those with whom they feel a genuine similarity (or to themselves alone). Defenders of an equality rooted in basic needs will push for universal programs that redistribute burdens

⁴⁸⁷ Ronald Dworkin subjects this argument from "wealth-talent" to a strong critique in *Sovereign Virtue*, 325-7.

among as wide a group as possible. The political success of social insurance programs will continue to depend on their ability to resist both extremes.

The fate of the Affordable Care Act, popularly known as Obamacare, helps to illustrate this point. Since its enactment in 2010, the law has sharply divided the American public, with roughly equal percentages of most demographic and economic groups steadily favoring and opposing it.⁴⁸⁸ A majority of Americans, including Republicans, Democrats and independents, tend to support some of its provisions, including its subsidies to lower-income citizens to buy insurance on the individual market and its prohibition on denying coverage to those with preexisting conditions.⁴⁸⁹ Yet one provision in particular, the individual mandate, which requires nearly all Americans to obtain health insurance or pay a penalty, has proven consistently unpopular since the law's enactment.⁴⁹⁰ The ACA's architects emphasize that the mandate is necessary for the coverage of preexisting conditions, since there must be enough healthy people in the risk pool to offset all of the sick ones. And according to the numbers, this should be a winning argument: Over half of all adults in the United States report that they or someone else in their household has a preexisting

⁴⁸⁸ The exceptions are Democrats and Republicans, large majorities of whom support and oppose the law, respectively, and ethnic groups, with large majorities of blacks and Hispanics favoring it, and smaller majorities of whites opposing. This data comes from the widely respected Kaiser Health Tracking Poll: <http://kff.org/interactive/kaiser-health-tracking-poll-the-publics-views-on-the-aca/#?response=Favorable--Unfavorable&aRange=all> (accessed February 2017).

⁴⁸⁹ Ashley Kirzinger, Elise Sugarman, and Mollyann Brodie, *Kaiser Health Tracking Poll: November 2016*, December 1, 2016, <http://kff.org/health-costs/poll-finding/kaiser-health-tracking-poll-november-2016/>.

⁴⁹⁰ A poll conducted in the wake of the 2010 mid-term election found that 88 percent of Republican voters and nearly half of Democrats—majorities of whom favored the law's other provisions—supported repealing the mandate. Kaiser Family Foundation, *Kaiser Health Tracking Poll: November 2010*, November 20, 2010, <https://kaiserfamilyfoundation.files.wordpress.com/2013/01/8120-f.pdf>. Democratic support for the mandate remained tepid in the ensuing years, while the intensity of opposition to it increased. See Kaiser Family Foundation, *Data Note: A Snapshot of Public Opinion on the Individual Mandate*, March 2012, <https://kaiserfamilyfoundation.files.wordpress.com/2013/01/8296.pdf>. By the end of 2014, it remained the least popular of the act's provisions, with 64 percent of the general public holding an unfavorable view. Bianca DiJulio, Jamie Firth, and Mollyann Brodie, *Kaiser Health Policy Tracking Poll: December 2014*, Dec. 18, 2014, <http://kff.org/health-reform/poll-finding/kaiser-health-policy-tracking-poll-december-2014/>.

condition according to the ACA’s definition.⁴⁹¹ These are therefore people who do not even have to imagine themselves behind a veil of ignorance to support social insurance, since they already know that they are the ones who would benefit from it. Add to these the many who may reasonably anticipate getting sick and struggling thereafter to find coverage, and one should have a strong contingency in support of compulsory risk pooling. Nevertheless, most Americans continue to resist the idea, and one of the very first acts that President Donald Trump took in office was to undermine it.

How are we to understand this phenomenon? Among the top objections that mandate opponents give are that they do not want the government dictating their choices and that the requirement to purchase insurance places too heavy a financial burden on them.⁴⁹² In other words, the mandate is at odds with personal responsibility, in that it does away with the choice of whether to insure or not, and it amounts to an unfair deal, particularly for those who are struggling economically but are forced to purchase expensive plans without a subsidy.⁴⁹³

The 2016 presidential election brought this view to the fore. Although healthcare was not a top-five issue for voters in either party, it almost certainly played a role in the election of Trump, who ran on a platform of repealing the ACA.⁴⁹⁴ Post-election polling data showed that 81 percent of Trump voters felt that Obamacare “went too far,” while half of Clinton voters felt the law did not

⁴⁹¹ Ashley Kirzinger, Bryan Wu, and Mollyann Brodie, *Kaiser Health Tracking Poll: Health Care Priorities for 2017*, Jan. 6, 2017, <http://kff.org/health-costs/poll-finding/kaiser-health-tracking-poll-health-care-priorities-for-2017/>.

⁴⁹² Kaiser Family Foundation, *Data Note: A Snapshot of Public Opinion on the Individual Mandate*.

⁴⁹³ Views of the mandate have been shown to be malleable in both directions. In particular, when respondents were reminded that requiring all Americans to have insurance could mean that people would be required to buy insurance they found too expensive or didn’t want, its favorability declined by 16 percentage points. Kaiser Family Foundation, *Data Note: A Snapshot of Public Opinion on the Individual Mandate*.

⁴⁹⁴ See, e.g., Olga Khazan, “How Obamacare helped Trump,” *Atlantic.com*, Nov. 9, 2016, <https://www.theatlantic.com/health/archive/2016/11/how-obamacare-helped-trump/507113/>.

go far enough.⁴⁹⁵ The economic composition of Trump’s support base helps to explain this. Half of all middle-class voters and nearly half of those making over \$100,000 a year supported the Republican candidate.⁴⁹⁶ These relatively wealthy Americans have been excluded from ACA subsidies, and as a result many found their health coverage increasingly unaffordable thanks to rising costs.⁴⁹⁷ By playing up the news of Obamacare premium hikes shortly before the November election, Trump tapped directly into their economic fears and resentments.

According to one pre-election poll, 66 percent of Trump’s supporters felt the American economy as a whole is rigged for people who receive government assistance.⁴⁹⁸ In particular, it appears that many working- or middle-class Americans resent that those earning slightly less than they do fare better in the insurance market thanks to ACA subsidies, while others who qualify for Medicaid seem to get decent healthcare for free.⁴⁹⁹ These voters would prefer a system in which they can buy lower-cost plans that serve their own needs, even if it means that less healthy purchasers have to pay more.⁵⁰⁰ It is worth recalling here that the ACA’s subsidies have proven popular over time, favored by the vast majority of Democrats and independents and at least six in

⁴⁹⁵ Paul Blake, “Election 2016 national exit poll results and analysis,” *ABC News Analysis Desk*, Nov. 9, 2016, <http://abcnews.go.com/Politics/election-2016-national-exit-poll-results-analysis/story?id=43368675>.

⁴⁹⁶ Jacquie Lee et al., “Exit polls by the numbers: Trump capitalizes on voter dissatisfaction,” *USAToday*, Nov. 9, 2016, <http://www.usatoday.com/story/news/politics/elections/2016/2016/11/09/exit-polls-numbers-trump-capitalizing-dissatisfaction/93500118/>.

⁴⁹⁷ Khazan, “How Obamacare helped Trump.”

⁴⁹⁸ Kai Ryssdal, “Poll finds Americans’ economic anxiety reaches new high,” *Marketplace*, October 16, 2016, <http://www.marketplace.org/2016/10/13/economy/americans-economic-anxiety-has-reached-new-high>.

⁴⁹⁹ Olga Khazan, “If not Obamacare, then what?” *Atlantic.com*, December 20, 2016, <https://www.theatlantic.com/health/archive/2016/12/if-not-obamacare-then-what/511130/> and Drew Altman, “The health care plan trump voters really want,” *New York Times*, Jan. 5, 2017, https://www.nytimes.com/2017/01/05/opinion/the-health-care-plan-trump-voters-really-want.html?_r=0.

⁵⁰⁰ Altman, “The health care plan trump voters really want.”

ten Republicans.⁵⁰¹ Such support reflects a view of insurance as a collective response to shared needs, actualized through a broad risk pool and assistance to those who require extra help to join in.⁵⁰² But the political fate of social insurance in liberal democracies depends on its also being perceived as broadly choiceworthy or fair for the individual, a view that in the ACA's case was undermined by its asking some people to pay substantially more for the same coverage than those they perceive to be their equals. If indeed any social welfare program that invokes the insurance principle will entail a balancing act between these two points of view, resistance to Obamacare came at least in part from its insufficient regard for the latter.⁵⁰³

Social insurance on this reading will not appeal to distributive perfectionists. Those who want a wholly collectivized response to basic needs will find the rubric of risk and its prudential calculations too constraining. Those who want reward to track desert, as defined at least in part by outcomes in a market economy, will find forced deviations from actuarialism and freedom of contract unfair. Yet as a regime, social insurance does pay deference to each view, and this is its central virtue.

Today, there are some signs that social insurance may be coming back into favor. Whereas mid-20th century theorists on both the left and the right decried the practice for obfuscating the true

⁵⁰¹ Kirzinger et. al, *Kaiser Health Tracking Poll: November 2016*.

⁵⁰² Indeed, some working-class Trump voters objected to Republican proposals to replace the ACA with personal savings accounts and tax credits on the grounds that these are “not insurance at all.” Drew Altman, “The health care plan trump voters really want.”

⁵⁰³ Dworkin offers a version of this same argument in his critique of Rawls's difference principle: “...[I]t seems callous to say that the only people for whom a theory of justice has concern are those whose lives are the most damaged, even though others, who work as hard as they can, are also seriously injured. Rawls has never been able satisfactory to explain why the members of his original position, ignorant of their own future status, would choose the difference principle out of their own self-interest, and the politics of the welfare problem show the practical shadow of that theoretical failure. Politicians who preach fairness to the ‘hard-working middle classes’...are also giving voice to a widespread instinct of justice.” Dworkin, *Sovereign Virtue*, 331.

aims of welfare, contemporary voices from the same corners now tout its advantages.⁵⁰⁴ This may be a salutary development, particularly if its intent is to find common ground in a deeply polarized political culture. If insurance is also to remain the central tool of welfare policy, however, it may be necessary to moderate our understanding of what a welfare state can do. By itself, social insurance will not eliminate poverty or rectify all unfair disadvantages, and it certainly will not reduce central government to a “night-watchman” role. Yet in its proper proportion, it may be uniquely capable of bringing about a roughly fair and stable polity: one in which the need for security is satisfied enough that citizens can concern themselves with higher pursuits, and in which the demand for distinction is paid enough deference that they will feel the desire to do so.

⁵⁰⁴ See, e.g., Anderson, “Common Property,” and Greg Weiner, “A Constitutional Welfare State,” *National Affairs* 29 (2016).