# An Essay on The Existence and Causes of Path Dependence

Scott E Page*

June 20, 2005

*I eat my peas with honey. I've done it all my life. It makes 'em taste quite funny, but it keeps them on the knife.* – an old Bostonian jump roping rhyme.

## 1 Introduction

That past events and choices can influence and in some cases determine the outcomes of current political, economic, and social processes is indisputable, but the extent to which history matters and the mechanisms through which it operates in particular contexts are more difficult to unpack. We can retell a particular history and explain how its many parts fit together. We can thus infer individual causal relationships—this appears to have happened because that happened—but we lack a framework for classifying and verifying the various ways in which history can matter.

An advantage of a framework is that it could be linked to empirical analysis. We could discern whether statistical evidence supported or refuted a particular claim of the extent and scope of the sway of the past. That said, empirical testing of a framework of causality is far from the only reason for constructing models of historical forces. Models discipline thicker, descriptive accounts (Gaddis 2002). They can help us understand the hows and whys. By boiling effects down to fundamental causes, they tell us where to look, and where not to look. They help us to identify conditions that are necessary and/or sufficient for past choices and outcomes to influence the present.

*spage@umich.edu. Professor Complex Systems and Political Science, CSCS Randall Labs, The University of Michigan, Ann Arbor, 48104. The author would like to thank the MacArthur Foundation and the James McDonnell Foundation for financial support. Aaron Bramson, Jenna Bednar, Anna Grzymala-Busse, and Ken Kollman made extensive comments on earlier versions. This paper also benefitted from the author's discussions with John Jackson, Skip Lupia, John Miller, Burt Monroe, Jim Morrow, Troy Tassier, Tarah Wheeler, and Ken Zick. Finally, participants at the two EITM summer workshops and the Santa Fe Institute Economics Workshop are thanked for their feedback.

The value of models makes the he lack of a theoretical framework to underpin the idea of historical dependence surprising. Compare the concept of path dependence to say that of Nash Equilibrium. We have refined the concept of equilibrium so that we now have quantile response equilibria, universally divine equilibria, equilibria in evolutionary stable strategies, and equilibria that satisfy the intuitive criterion. Some may interpret these visions and revisions as splitting the hairs on the heads of so many angels dancing on pins. I disagree. These many refinements clarify our thinking about what it means to be rational to levels that could not have been imagined much less attained without the game theory framework.

Models and frameworks serve powerful functions even just as analogies. Consider the impact of the idea that increasing returns (increases in benefits from some choice or action as more people take that action) can generate path dependence. (David 1985). Path dependence in its loosest sense means that current and future states, actions, or decisions depend upon the path of previous states, actions, or decisions. Of late, path dependence has become a popular conveyor of the more general idea that the timing of events matters (Pierson 2004) and as way to model institutional stickiness (Crouch and Farrell 2004).

History dependence and path dependence are real phenomena. Theoretical, historical, and empirical studies of path dependence run the gamut, covering topics ranging from the selection of institutions (North 1991), to the formation of government policies (Hacker 2002), to the choice of technologies (David 1985, Arthur 1994), to the location of cities (Arthur 1994, Page 1998), to pest control strategies (Cowan and Gunby 1996), to the formation of languages and law (Hathaway 2001). The existence of path dependence appears indisputable. But are all of these cases the same? In each of these cases does history operate in the same way? Assuredly not. The micro-level processes that produce path dependent government policies, pest control strategies, and laws differ markedly in their levels of inertia. A law is much more strongly and differently connected to past laws than a pest control strategy is to past pest control strategies. Further, historical forces constrain laws to be similar to past laws. The opposite would seem to be true for pest control strategies. The optimistically named exterminator wants to kill precisely those pests that previous strategies didn't kill. As these two examples show, not all path dependent processes work the same way. Our models should not treat them as if they do. Better yet, we should have the flexibility when looking at history to choose from a box of lenses.

This is not just a matter of theorists having precise definitions. The stakes here are large. Path dependence may explain why some countries succeed and others do not (Easterly 2001). Economic growth models predict that less developed countries should catch up with their richer counterparts, but that has not happened. Doug North has shown how country level success depends upon the proper build up of institutions, behaviors, and law. Some countries start out eating their peas with honey and never move to the fork or spoon.[1] Historically, it has been shown that gradually moving to that fork or spoon isn't easy. Economies often have to take big jumps from more backwardness to status quo technologies (Gerschenkron 1952).

To help make sense of the weight and sway of history, I describe here a general framework within which we can formally define various forms of history dependence. I differentiate between *path dependence*, where the path of previous outcomes matters, *state dependence* where the paths can be partitioned into a finite number of states which in contain all relevant information, and what I call *phat dependence* where the events in the path matter, but not their order. I also distinguish between early and recent path dependence, and perhaps most importantly, between processes in which *outcomes* are history dependent and those in which the *equilibria* depend upon history. By equilibria here I mean limiting distributions over outcomes. So a process that converges to a probability distribution over outcomes would attain an equilibrium according to this definition. If the equilibrium outcome can depend upon history, then there must necessarily be multiple equilibria.

These definitions have value in their own right. They replace sloppy, vague notions with sharp, precise terminology and notation. They have functional value as well. They allow us to expand the set of possible causes of history dependence, to identify necessary conditions for history dependence to occur and to clear up confusions about the relationship between history dependence and increasing returns. It is within this framework that I can show that increasing returns does not imply nor is it implied by path dependence.

I also present examples of these various forms of path dependence within a dynamical systems framework using two broad classes of models. I base the first on dynamical systems and the second on choice theory. These models help to reveal the causes of path dependence. The proximate cause of history mattering differs in the two classes of models. In the first, history has force. The past exerts sway over the present. A decision to prohibit women from voting effects how women see themselves in relation to men. When given the vote, women cannot escape all effects of their past denial of rights. In the second, the cause is more direct. A decision to provide social security for the aged lowers the economic and political costs to extending those benefits to orphans and the infirm. When written in the dark lead of mathematics, the line between forces and externalities appears crisp, but that is not so. Cognitive attachments are externalities in our heads (Medin and Atran 2004, Bednar and Page 2004), and externalities between choices change the incentives for making subsequent choices. Changes in incentives can be equivalent to changes in force. Both classes of models are abstract. Balls and urns, in particular, bear little resemblance to people and institutions. That said, I make every effort to connect these models to the real world of politics, or at least to interpretations of that real world.

[1] I use a spoon for my peas. It's not socially acceptable but coming from a big family, I had to eat quickly.

I have divided the remainder of the essay into four parts. In the first part, I describe what I see as the limitations in our current understanding of path dependence. In the second part, I provide a partial taxonomy for the ways that history can matter. I distinguish between state dependence, phat dependence, and path dependence as well as between early path dependence and recent path dependence. I begin by making an important distinction between path dependent outcomes and path dependent equilibria. I supplement these definitions with examples using the ball and urn models. I also construct examples to show that increasing returns is neither necessary nor sufficient for historical dependence (David, 1985, North 1991, Arthur 1994, and Pierson 2000). This lack of necessity and sufficiency implies that we must seek elsewhere when looking for path dependence, and that evidence of increasing returns is not enough. We need more.

In the third part, I construct a second class of models that relies on externalities between actions or outcomes. Within that class I show that increasing returns cannot cause phat or path dependence, and I show that negative externalities are necessary. I also show that externalities whose values do not change over time cannot cause path dependence. They can only cause phat dependence. Thus, evidence of path dependence requires a build-up of behavioral routines, social connections, or cognitive structures around an institution. This micro-level stickiness is implicit in historical accounts like those that North (1990), Pierson (2004), and Hacker (2002). It has even been made formal by Bednar and Page (2005) in a model of institutional and game ensembles. Thus, the historical and theoretical evidence strongly suggests that while externalities can cause phat dependence, those externalities must change over time to generate path dependence. In the final section, I briefly discuss the relationship between path dependence and rationality.

# 2 Common Misunderstandings

A survey of the literature on path dependence reveals an emphasis on four related causes: increasing returns, self reinforcement, positive feedbacks, and lock-in. There are subtle differences between these four. Increasing returns means that the more a choice is made or an action is taken, the greater its benefits. Self reinforcement means that making a choice or taking an action puts in place a set of forces or complementary institutions that encourage that choice to be sustained. With positive feedbacks, an action or choice creates positive externalities with that same choice if made by other people. Positive feedbacks create something like increasing returns, but mathematically, they differ. We might think of increasing returns as costs or benefits that rise smoothly as more people make a particular choice and of positive feedbacks as little bonuses given to people who already made that choice and who will make that choice in the future. Finally, lock-in means that one choice or action becomes better than any other one just because everyone else has made that choice. For better or worse, the English language is locked-in within corporate and legal America.

Paul David's example of the QWERTY typewriter exhibits all of these features. There were increasing returns to buying a QWERTY typewriter because as more were sold, the costs of marketing QWERTY's fell. The QWERTY was also self reinforcing. Textbooks for learning how to type were all based on the QWERTY layout. The QWERTY also created positive feedbacks. One QWERTY typist could type on another person's QWERTY. Eventually, once there were enough QWERTYs in use, it became locked in as a technology. We are all part of QWERTY nation.

Before describing a framework for thinking about path dependence, I want to highlight five common misunderstandings that seem to exist in our heads as well as in the literature. The first misunderstanding is the conflation with path dependence and increasing returns. This may be the most provocative result stated in this essay. Neither increasing returns nor path dependence implies the other. They are logically distinct concepts. Increasing returns are neither necessary nor sufficient for path dependence. Nevertheless, the conflation of increasing returns with path dependence is understandable. A process that generates multiple equilibria that depend upon the path taken would seem to require that the more an outcome occurred in the past, the more likely it is to occur in the future. That logic is flawed, however; it ignores the possibility of more complicated interactions between actions.

The first misunderstanding results in a narrowness of focus. John Von Neumann is said to have once referred to the term nonlinear functions as equivalent to the term non-elephant animals. The same sentiment might be expressed here. Increasing returns can create path dependence but so can almost any type of negative externalities. These could be caused by any kind of constraint: spatial, budgetary, or even cognitive. The set of all externality classes is so large that it is difficult to characterize. We need to explore more of the ark. So far, we've only been considering the elephant of increasing returns. By thinking too simply, we've ignored a whole world of possibilities for path dependence.

Consider Hacker's (2002) analysis of health care policy. Roosevelt pushed for and got pension insurance, social security, but not national health care. The burden of providing health care fell to others. Firms might have emerged that provided health care. We might then buy health insurance just as we now by home or care insurance. Instead, employers began to provide health insurance. Why they did is complicated. During World War II, wage controls prevented firms from increasing pay, but they could increase benefits. Employer provided health benefits can then be seen as a result of market forces. Once some employers began to provide health care, the labor movement and worker expectations led to more and more employers providing health benefits to the point where now, health benefits are expected.

The history can be see as increasing returns with respect to employer provided health benefits, but that is an oversimplification. Hacker's full story is one of externalities accumulating on top of externalities. This is a *negative externality* imposed on business.

The omission of benefits from the wage controls created a positive externality for businesses that wanted to provide higher wages. The expectations that employers provide health insurance created a negative externality on employers. The build up of firms that catered to firm and not individuals created positive externalities with employer provided benefits and negative externalities with individual health care plans. No one disputes that history mattered, but we would like to be able to make sense of the causes. We would like to know the relative weight of the negative externalities with respect to the positive externalities.

The second misunderstanding involves a misassignment of credit. In many of the examples of path dependence, while increasing returns do exist, negative externalities are the true cause. This is not merely a reframing of positive relative returns as negative relative returns. It requires a fundamental rethinking of what causes path dependence. Any constraint, be it a budget constraint, a spatial constraint, or a time constraint, imposes negative externalities and can create path dependence. The logic of constraints applies to competing technologies, legal doctrines, and city locations. In each case, the exclusion of other options drives the path dependence. This is even true in cases where increasing returns are extremely powerful such as in the build up of offensive forces and weaponry (Van Evera 1998). Near the end of the paper, I walk through this logic in some detail using Paul David's QWERTY example.

The third misunderstanding results from not looking closely enough at a particular model. The most commonly invoked model of path dependence is the Polya Process. In the Polya Process, the set of past events, but not their order, determines the probability distribution over outcomes. Let me say this more directly: *in the Polya Process, the order of history does not matter*. More formally: *the Polya Process is* phat *dependent but not* path *dependent.* According to the logic of the Polya Process, if we want to know why Lenin, Washington, or Tubman took a particular action we could consider his or her past in any order. We could read the pages of Washington's biography in random order and the explanation for why he freed his slaves at the end of his life would be just as complete. My limited reading of history suggests that while historians do not believe that all history is relevant, they often mean much more than phat dependence. They seem to care a lot about the sequencing of events.

The fourth misunderstanding results from a failure to disentangle individual outcomes and long run equilibria. Path dependence can be defined with respect to either. Outcome path dependence implies that what happens in this period depends upon the past. This does not mean that what happens in the long run depends upon the past. The long run may well converge to a unique equilibrium. Equilibrium path dependence implies that what happens in the long run depends upon the path. Clearly, it is the latter that people mean when they say that history matters, but the evidence that we can muster is often only of former type. We can show that what happened now depended upon happened then. A system can exhibit outcome dependence yet still have a unique equilibrium.

The final misunderstanding concerns a conflation between what might be called *early path dependence* and *sensitivity to initial conditions*. Sensitivity to initial conditions, a term borrowed from chaos theory, refers to deterministic dynamical systems in which the trajectory or the equilibrium depends sensitively on the initial point of the system. Extreme sensitivity to initial conditions implies that small changes in the initial point lead to divergent paths. Minor initial changes have enormous implications. These systems are deterministic. They are often recursive. A single function $f$ is recursively applied to an outcome, and small differences get magnified with each recursive iteration. Early path dependence describes processes in which early random outcomes shape the probability distribution over future histories. They do not determine it. They shape it. For example, in the evolution of common law, the doctrine of *stare decisis et non quieta movere*, states that past decisions should stand, but this is not a hard and fast rule (Hathaway 2001). Decisions can, and do, get overturned as can the logic that underpins them. The doctrine of *stare decisis* implies that weight be placed upon the early history of outcomes but it does not imply sensitivity to initial conditions. The future is not deterministic, but stochastic and biased toward early decisions. Or, to return to Hacker's analysis, the New Deal did not set in motion a deterministic process. What it did was relieve corporations of some of the burden of providing retirement benefits and laden them with the responsibility for health benefits. If we think through the counter factual, had FDR gone after health benefits, most of us would agree that the political landscape would look differently.

# 3 Dynamic Processes

I begin with a description of a dynamic process that produces *outcomes* at discrete time intervals which I index by the integers, $t = 1, 2, ....$[2] I denote the outcome at time $t$ as $x_t$. In a more general model, in each period in addition to the outcome there may also be other information, opportunities, or events that arise in a given time period. These can be described as the *environment at time t*. This contains exogenous factors that influence outcomes. A *history at time T*, $h_T$ is the combination of all outcomes $x_t$ up through time $T-1$ and all other factors, the $y_t$, up through time $T$. In the ball and urn models that I describe in this section, there will be no $y_t$ terms, so the history consists only of past outcomes.

A dynamic process also has an *outcome function G* that maps the current history into the next outcome. The outcome generated by a dynamic process can then be written as follows:

$$x_{t+1} = G_(h_t)$$

[2]Notice that this is a discrete time process. I could also write the process as occurring in continuous time. Discrete time processes are far easier to analyze.

The outcome function can change over time so it is indexed by $t$. Not also that the function $G_t$ is not necessarily deterministic. It can also generate a probability distribution over outcomes. This will be the case in many of the examples that I consider. History dependence need not imply deterministic dependence. It need only imply a shift in the probabilities of outcomes as a function of the past.

## 3.1 Outcome and Equilibrium Dependence

This stark framework enables me to distinguish between two ways that history can matter. History can matter in determining the outcome at time $t$, $x_t$. I call this *outcome dependence.*

*A process is* **outcome dependent** *if the outcome in a period depends upon past outcomes or upon the time period.*

History can also matter for the limiting distribution over outcomes. I call this *equilibrium dependence.* It is possible, in fact perhaps likely, that a process might not attain an equilibrium distribution over outcomes, but it does in all but one of the processes that I consider here.

*A process is* **equilibrium dependent** *if the long run distribution over outcomes depends upon past outcomes.*

Notice that I am adopting an expansive notion of equilibrium, that of the long run distribution of outcomes converging. Alternatively, I might have required that the outcome functions $G_t$ converge in each period to a common distribution over outcomes. Were I do to that I would rule out process that converge to equilibrium cycles or patterns. I return to this point later.

Two other observations merit mention. First, as touched upon earlier, equilibrium dependence requires multiple equilibria. If the process generates a unique equilibrium, then there can be no influence of history on the equilibrium outcome. Second, equilibrium dependence implies outcome dependence. If the equilibrium distribution over outcomes depends upon the past, then so must the outcomes in individual time periods. If history determines which technology will be used in the long run, then it must also determine which technology is selected in the periods that make up the long run. The causality does not hold in the opposite direction. History can matter for the outcome in every period, but not matter for the distribution over outcomes. This subtle point is one of the common misunderstandings that I described in the previous section. I therefore state it as an observation.

**Observation 1** *History can matter at every moment in time, but not matter in the long run.*

Thus, whether history matters in some real sense depends upon what we want to explain. Two historical examples help clarify the distinction and highlight why we might or might not care about outcome dependence. Ferejohn (1991) describes electoral competition in Early Stuart England. At that time, people, okay men, alternated holding elected office. There was no meaningful electoral competition. The outcome in any one period depended upon the past: the person who last held the seat no longer held the seat. Outcomes were therefore history dependent. However, the equilibrium distribution of who held the seat did not depend upon history. Each eligible person held the seat roughly an equal number of times. The equilibrium distribution over outcomes does not depend at all on who held office in any one term.

To give a more weighty example, the term "manifest destiny" introduced by U.S. politicians in the 1840's implied an unstoppable processes of continental expansion. Let us suppose that there was such a thing as manifest destiny. At any moment in time, the path of Western expansion depended on the path taken: whether farmers, ranchers, or trappers led the way, or whether the Northern or Southern part of the continent developed first. The process of expansion generated history dependent outcomes: the regions that gained in population in 1852 where dependent upon where people moved in 1849, 1850, and 1851. But, the eventual outcome may not have depended on history. Had not been found at Sutter's mill, California might probably still be our most populous state.

That is not to say that equilibrium dependence is rare. It is not. Most of the cases discussed by Pierson (2004) are examples of equilibrium dependence. In his excellent analysis of the creeping accumulation of judicial power, he is making clear that what happens now *and* what happens in the future are dependent on what happened in the past. The distribution over outcomes, not just this one outcome, have changed.

The same might not be said of journalistic accounts of the shifts in balances of power within our federal system from those between the branches of government to those between the states and the federal government. These ongoing relationships experience ebbs and flows but a robust federation requires that they not shift too far out of balance (Bednar 2005). Individual outcomes may be path dependent, but Madison certainly hoped that in the long run, any gains in power by one branch of government would create balancing incentives among the other branches.

## 3.2 The Ball and Urn Models

To describe the specific ways that history can matter either for outcomes or for equilibria, I construct examples using the class of ball and urn models. These models generalize the familiar Polya Process. Though they rely only on balls of different colors being drawn from a single urn, they prove capable a surprising and fortuitous breadth of application. The breadth is surprising in that as I parse the concept of path dependence many ways and create a taxonomy of types of path dependence, I find that I can construct a simple ball an urn example for each type. This is fortuitous in that the examples clarify the definitions. We need not abandon the simplicity of the urn model to reproduce the necessary definitional breadth.

The ball and urn model assumes a collection of various colored balls in an urn. In each period, a ball is selected from the urn, and depending on the color of that ball selected, other balls may be added or removed from the urn. The selection of the ball plays the role of the *outcome function.* Because the ball is selected randomly, the probability of an outcome depends upon the composition of the urn: how many balls of each color it contains. In almost all of the examples, I assume two colors of balls: maroon, which I denote by $M$ and brown, which I denote by $B$. To provide some real world context, rather than think of balls of various colors being drawn from an urn, you can think of a society adding new institutions. These institutions can be either market based ($M$) or bureaucratic ($B$). Colors and institutional choices can be used interchangeably. An outcome $M$ can be thought of as a maroon outcome or as a market outcome.

Given this setup, a history of outcomes can be written as a sequence of $M$'s and $B$'s. I first describe processes in which outcomes do not depend on time or on these histories of outcomes. Such processes are called *independent.*

*A process is* **independent***, if the outcome in any period does not depend upon past outcomes or upon the time period. An independent process can be written as follows:*

$$x_{t+1} = G(\cdot)$$

My first example is a Bernoulli Process. In a Bernoulli Process, no new balls are ever added to the urn, so the probabilities of selecting a maroon or brown ball never change.

**Example 1: A Bernoulli Process** The urn contains $M$ maroon balls and $B$ brown balls. Each period a ball is chosen randomly and then put back in the urn. The probability of drawing a maroon ball equals $\frac{M}{(M+B)}$ and the probability of drawing a brown ball equals $\frac{B}{(M+B)}$ in every period.

In a Bernoulli process, every period is identical, so the process is independent.[3] Many canonical random processes are assumed to be independent: the outcome of a coin flip or the roll of a die, or the sex of a child. Each outcome in a sequence is independent of the previous outcomes. Examples of independent processes from the political and economic world are harder to come by, as we typically think that the past matters in some way.

[3]Interestingly, given the human tendency to predict and recognize patterns in random sequences, people often see path dependency when none exists such as trends in stock market returns or hot streaks by basketball players (Gilovich, Vallone,and Tversky,1985).

## 3.3 History Dependent Outcomes and History Dependent Equilibria

A process that is not independent can be *history dependent*: the current outcome or both the current outcome and the equilibrium distribution over outcomes could depend upon the past history of outcomes. In either case, there remains the question of how much and to what extent history matters for outcomes and equilibria.

I distinguish between three types of history dependence: *state dependence*, *phat dependence*, and *path dependence*. These types can be thought of as levels of history dependence with state dependent processes being the least and path dependent processes being the most history dependent.

### 3.3.1 State Dependence

In some cases, it is possible to partition the space of all histories into a finite number of sets: $\{s_1, ...s_N\}$ such that the outcome function at each moment in time depends only on the set to which the current history belongs. These sets are then called *states.* In the ball and urn models, the number of maroon and brown balls contained in the urn can represent the *state* of the process but only if the number of possible combinations of balls is finite. If there can be infinitely many balls in the urn, then in this class of models, there is no state, per se.

Suppose, for example, that two political parties alternate top position on the ballot from election to election. The outcomes in period $t$ would be which party was on the top of the ballot. To know the outcome at time $t+1$ all that needs be known is the outcome at time $t$, therefore all histories can be partitioned into two states: those in which the first party was last at the top of the ballot and those in which the second party was at the top of the ballot. Or, suppose that the federal reserve board were to use a fixed rule for setting the prime rate that depended only on the current inflation rate. If so, the prime rate (the outcome) would depend upon history only in so far as the current inflation rate depended upon history. History matters, but any two histories that result in the same inflation rate (the same state) are equivalent with respect outcomes.

Given that there exists only a finite number of states, it is possible then to write a mapping from each history into one of these $N$ states. It remains to describe how the states change over time. There is assumed to exist a *state transition rule*, $T$ that maps the current state $s_t$ and (possibly) the current outcome $x_t$, into the next period's state. This can be written as $s_{t+1} = T(s_t, x_t)$. The state transition rule can be random or deterministic, but it cannot depend upon the entire history. It can only depend upon the finite states.

*A process is* **state dependent***, if the outcome in any period period depends only upon the state of the process at that time. A state dependent process can be written as follows:*

$$x_{t+1} = G(s_t) \quad \text{where} \;\; s_{t+1} = T(s_t, x_t)$$

Since the outcome only depends upon the state, this implies that $G_t = G$ for all time periods $t$. Such processes are commonly called *Markov Processes.* I refer to them here as state dependent processes to highlight their differences with phat dependent and path dependent processes. For obvious reasons, these processes generate history dependent outcomes. The history determines the state and the state in turn determines the distribution over outcomes.

I now add two further restrictions. A state dependent process is said to be *stationary* if the state transition rule $T$ is the same in every time period. A state dependent process is said to be *ergodic* if through some series of states it is possible to get from any one state to any other. I can now state one of the most important, and I dare say neglected theorems, in historical analysis, the Ergodic Theorem.

**The Ergodic Theorem:***A stationary, ergodic state dependent process generates a unique equilibrium distribution over outcomes.*

The Ergodic Theorem raises the bar quite high for anyone who wants to claim that history matters in the long run. The theorem says that (i) if it is possible to define a finite set of relevant states that determine the next outcome, (ii) if the mapping from states to states as a function of outcomes does not change, and (iii) if it is possible (no matter how unlikely) to get from any state to any other, then the process converges to a unique distribution over outcomes. That equilibrium, being the only one , is also stable. The phrase "E.T. Phone Home" can a helpful reminder here. The Ergodic Theorem (E.T) states that the distribution is called to a single place (home).

The Ergodic Theorem does not deny outcome dependence. In fact, only the most trivial state dependent processes do not exhibit outcome dependence. It just says that if we were to run a process many times and if each time we were to bin the outcomes and create a distribution, we would find that the distributions were all the same. In the long run, the history of outcomes wouldn't have mattered. This leads to my second observation, which is a restatement, albeit an important one, of the Ergodic Theorem.

**Observation 2** *Equilibrium dependence requires a changing process, a big space, or divergent paths.*

This observation can be restated more formally. To have multiple equilibria, one of the four core assumptions of the Ergodic Theorem must be violated: (i) the outcome function must change over time (ii) the transition function between states must change over time or (iii) the number of states must not be finite or (iv) getting from some states to some other states must be impossible via any sequence of states. To assume the first is to say that the world changes over time. If it changed as a function of an outcome, then this would, in effect, assume path dependence. That's perfectly admissible. One interpretation of the concept of *lever points* is that they are moments in time in which an outcome changes the course of history, and in doing so changes the outcome function. If that happens, then certainly history would matter.

A process might also violate stationarity. The state transition function could change. For example, if the state represented a voter's political ideology: conservative, liberal, or independent, the transition probabilities between ideologies could be a function of the ideologies of other voters. If so, the transition probabilities would change over time and multiple equilibria could be sustained.(Page 2005).

It is also possible to violate the assumption of a finite number of states. Even in simple ball and urn models that assumption can fail to hold so there is no reason to believe it would hold in the real world of people and institutions. And, even if it does, if the number of states is huge, it could be that the equilibrium predicted by the ergodic theorem, which is an equilibrium over the probabilities over outcomes might not be as informative as looking at the recent path. The U.S. Mint recently allowed each of the fifty states to issue a new quarter. The states were placed in the order that they were admitted to the Union with New Hampshire going first and Hawaii last. Suppose that this process then repeats, with New Hampshire issuing a new quarter. Each U.S. state could then be represented by a state in the process. The equilibrium of this process is that each state (however conceived) is chosen with equal probability.

Finally, ergodicity can be violated. Sometimes, two roads really do diverge in a yellow wood, and when we take one we cannot go back and take the other just as fair.[4] Making the case that a process is not ergodic is not as easy as it sounds. It has to be shown that some state can never be reached by *any path* from some other state.

### 3.3.2 Phat and Path Dependence

Phat and path dependent processes all histories can generate equilibrium dependence as well as outcome dependence. The distinction between state dependence, on the one hand, and phat and path dependence, on the other, is more subtle. A history of past outcomes and opportunities can always be written as a state. Doing so transforms any path or phat dependent process into a state dependent process. This would seem to suggest that all processes are state dependent. However, the number of histories and sets of histories could be infinite, and the set of states is assumed to be finite. Therefore, it is not possible to reinterpret paths and sets of paths with states.

If the history of outcomes matters but not the order in which they occurred, I define

[4]Had Frost had more of a mathematical bent rather than writing, "Yet knowing how way leads on to way, I doubted if I should ever come back", he would have written "Yet knowing the paths were not ergodic, I knew that I should never come back."

the process as *phat dependent*. The word *phat* describes the former type of processes in two ways. It reorders the letters in path, demonstrates that the order does not matter, and it echoes the importance of thickness in our descriptions.

*A process is* **phat dependent** *if the outcome in any period depends upon the set of outcomes and opportunities that arose in a history but not upon their order. A phat dependent process can be written as follows:*

$$x_{t+1} = G_t(\{h_t\})$$

Where $\{h_t\}$ denotes the set of outcomes up to time $t$.
If the order of the history of outcomes also matters, I define the process as *path dependence*. The definition of path dependence looks almost identical to that of phat dependence. The lone difference is that for a path dependent process the outcome function depends upon the vector of history $h_t$. Changing the order of $x_1$ and $x_2$ could change the outcome produced by $G_t$.

*A process is* **path dependent** *if the outcome in any period depends upon history and can depend upon their order. A path dependent process can be written as follows:*

$$x_{t+1} = G_t(h_t)$$

To put this in the context of the ball and urn models, the *path* is the ordered set of all previous outcomes. The paths $MBM$, $BMM$, and $MMB$ create the same set of outcomes, namely $\{M, M, B\}$. If only this set matters and not the order in which the outcomes arose, the process is *phat dependence*. For example, jurors voting on a defendant's guilt or innocence often sequentially reveal their opinions. A juror's opinion might change in response to the expressed opinions of others. The process would be phat dependent if each juror considered only the number of previous jurors voting guilty and innocent when making her own opinion known. The process would be path dependent if each juror also took into account the order in which the other opinions were voiced. This distinction has many implications. Testing for phat dependence a different econometric model than testing for path dependence (Jackson 2004), a point I return to later. How we tell a history also depends upon the nature of the historical process. If a process is phat dependence, we can describe the history leading up to a point in any order, what came first second or third does not matter.

My second example is the Polya Process. In the Polya Process, the urn initially contains one brown and one maroon ball. In each period, a ball is selected and returned to the urn, and another ball is added to the urn of the same color as the selected ball. At least metaphorically, this process is thought to capture the phenomenon of increasing returns that Arthur (1994), David (1985), Pierson (2000) and others elaborate in the context of institutional choices. The more often that market solutions (maroon balls) are chosen, the more likely that they will be chosen in the future.

**Example 2: The Polya Process** Initially, $M = B = 1$. In any period, if a brown (resp.a maroon) ball is selected then it is put back in the urn together with an additional ball of the same color.

The Polya Process is equilibrium dependent. Not only can the process converge to more than one ratio of maroon and brown balls, it can converge to *any* ratio of maroon and brown balls. Depending upon the history of outcomes the urn could eventually contain 80% maroon balls and 20% brown balls or it could contain 63% maroon balls and 37% brown balls. Once there are enough balls in the urn, the current ratio remains constant, and balls continue to be selected in those proportions. [5] Thus, the Polya Process is equilibrium dependent, and therefore it also outcome dependent.

The Polya process is not *path dependent*. It is only *phat dependent*. The outcome at time $t$ only depends on set of past outcomes not on their order. In period five, if there are three additional maroon balls in the urn and one additional brown ball, the outcome is the same regardless of the order those balls were selected. It does not matter whether the three maroon balls were chosen first and then one brown ball or whether the one brown ball was chosen and then three maroon balls. In addition, both of these paths are equally likely. This is easily shown. The probability of the first path is $(1/2)*(2/3)*(3/4)*(1/5)$. The probability of the second path is $(1/2)*(1/3)*(2/4)*(3/5)$. This holds in general. For any distribution of balls, any history that generates it, is equally likely. So not only is that set all that matters for the future, the set does not tell us much about the past.

**Observation 3** *The Polya Process is phat dependent.*

Even though the Polya Process is only phat dependent, this does not imply that the real world situations it has been used to described do or do not depend upon the order of the path. Those are, of course, empirical questions. Evidence tilts strongly in favor of both types of dependence. Earlier I gave examples of path dependent processes, so let be describe some instances where we ignore order. In calculating the ideologies of congresspeople, the order of the votes within a congress is not considered, just the votes themselves.[6] A model that used ideology to predict an outcome - in this instance, a current vote, then that model would be assuming phat dependence. Second, in predicting the number of votes that a congressperson, let's say my congressperson, John Dingell, will get in his next congressional race, we regress an equation based his ideology, his reputation, his positions on the issues, his opponent, his supporters and his opponent's supporters, the amount of money he and his opponent have and characteristics of his

[5]The Polya process has many interesting features that lie beyond the scope of this analysis.

[6]Note: If the number of ideologies were finite, then this could be a state dependent process.

district. This is a thick description of the race, but it is not an ordered description. It is phat dependence not path dependence.

A next, and obvious question to ask is whether every phat dependent process is equilibrium phat dependent or whether a process could exhibit only outcome phat dependence. The next example, the Balancing Process, shows that outcome phat dependence only is in fact possible. The Balancing Process is not *independent.* Outcomes depend on the past. Nor is it *state dependent.* The number of states are infinite. Therefore, it is phat dependent. However, it does not generate equilibrium phat dependence. It has a unique equilibrium distribution.

In this process, the probability of selecting a ball of a given color again depends upon the colors of the previous balls chosen, but in this case the feedbacks are negative. Instead of inserting a ball that is the same color as the ball that was drawn as in the Polya Process, now a ball of the *opposite* color is added to the urn. This processes balances the outcomes. As more maroon balls are chosen brown balls become more prevalent in the urn and more likely to be chosen.

**Example 3: The Balancing Process** Initially $M = B = 1$. In any period, if a brown (resp.a maroon) ball is selected then it is put back in the urn together with an additional ball of the opposite color.

To see why the balancing process cannot generate multiple equilibria, suppose that the process converged to something other than an equal number of maroon and brown balls. Imagine an urn with a large number of balls 60% of which are maroon and 40% of which are brown. From that point onward, maroon balls would be more likely to be selected. Selecting these maroon balls would add brown balls to the urn, increasing the proportion of brown balls above 40%.

This balancing phenomena mimics political and economic processes that must satisfy competing interests. I previously discussed balance of powers in a robust federation, but this balancing can occur at other scales as well. If one constituency prefers market solutions and another prefers bureaucratic solutions, then successes by one constituency may result in the mustering of greater political forces by the other: The addition of another market based institution may create more pressure for future bureaucratic solutions. In the model this is captured by adding a brown ball to the urn whenever a maroon ball is selected and adding a maroon ball, whenever a brown ball is selected. Rotation schemes (Kollman 2003) in the European Union are an extreme example of a concern for fairness but more subtle balancing may occur in the selection of locations for political conventions, the World Cup, or the summer and winter Olympics. Votes by individuals may also balance interests if they want moderate policies.[7]

[7] I would like to thank participants at the NSF sponsored EITM at the University of Michigan for these examples.

## 3.4 Increasing Returns and Equilibrium Dependence

As I discussed at length earlier, much of the literature on path dependence emphasizes increasing returns. The connection between increasing returns and path and phat dependence is not as straightforward as it might seem. My next two examples help to clarify this relationship. The informal definition of increasing returns goes as follows: the more an outcome occurs, the higher the relative return to that outcome, and therefore, the more likely it occurs in the future. Increasing returns within the class of urn processes is a simpler concept. It leaves out the middle step in the causal chain. I present here a precise definition of increasing returns in the class of urn models.

*An dynamic process generates* **increasing returns** *if an outcome of any type in period $t$ increases the probability of generating that outcome in the next period.*

The Polya process satisfies increasing returns and exhibits equilibrium phat dependence, though not path dependence. It is thus possible for an urn process to exhibit increasing returns and to generate multiple equilibria. That does not mean that all processes with increasing returns generate multiple equilibria nor does it imply that all process that generate multiple equilibria satisfy increasing returns. There is no logical implication in either direction as I show in the next two examples. The first example relies on red ($R$) and green ($G$) balls as well as maroon and brown balls. I call it the Balancing Polya Process. It is a combination of the Balancing Process and the Polya Process.

**Example 4: Balancing Polya Process** Initially, $M = B = R = G = 1$. In any period, the ball selected is returned to the urn. In addition, if a red ball is selected, a maroon ball is added to the urn. If a maroon ball is selected, a red ball is added to the urn. If a green ball is selected, a brown ball is added to the urn. If a brown ball is selected, a green ball is added to the urn.

It is straightforward to show that this process exhibits equilibrium phat dependence. To see this, paint the red balls maroon and the green balls brown. Doing so creates the Polya Process which as we know is equilibrium phat dependent. But, the Balancing Polya Process does not satisfy increasing returns. In any given period, choosing any color ball decreases the probability of picking that ball in the next period. Thus, increasing returns are not necessary for equilibrium dependence. [8]

The Balancing Polya Process provides a hint as to how complementarities between

[8] The Balancing Polya Process is not the only example that demonstrates a lack of necessity. They are remarkably easy to construct. The *Locked Out Process* is defined as follows. Initially, $M = B = 10$. In the first nineteen periods, a ball is selected and removed from the urn. In subsequent periods, the one remaining ball is repeatedly selected and returned to the urn. The Locked Out Process exhibits equilibrium dependence. It converges to one of two equilibria. Eventually, either all brown or all maroon balls are selected.

outcomes – red outcomes creating an environment favorable for maroon outcomes in the future – can generate equilibrium dependence. Many scholars, North (1990), Pierson (2004), Ikenberry (2001) and Bednar (2005) argue for strong complementarities between institutions. This examples provides some intuition for how those complementarities, and not increasing returns per se, can generate equilibrium dependence.

The next example proves a lack of sufficiency. It exhibits increasing returns for both maroon and brown balls but does not generate path, or even phat dependent equilibria. I call this the Biased Polya Process, as the brown balls have an advantage.

**Example 5: The Biased Polya Process** Initially $M = 1$ and $B = 2$. In each period, a ball is selected. If a maroon ball is selected, it is put back in the urn together with another maroon ball and another brown ball. If a brown ball is selected in period $t$ is put back in the urn together with $2t$ additional brown balls.

It can be shown that a brown ball is selected with probability one. Once that brown ball is selected, the probability that the next ball is brown exceeds 75%. Eventually, the proportion of brown balls in the urn converges to 100%, so this process generates a unique equilibrium. Selecting a brown ball clearly satisfies increasing returns. Select a brown ball and a brown ball is more likely to be selected in the next period. Surprisingly, maroon balls also satisfy increasing returns. Select a maroon ball, and the probability of selecting a maroon ball in the next period also increases.[9]

**Observation 4** *Increasing returns are neither necessary nor sufficient for equilibrium dependence.*

In light of this observation, why the common conflation between increasing returns and path dependence? The logic goes as follows: For a process to exhibit equilibrium dependence it must converge to at least two distinct equilibrium probability distributions. For this to happen in an urn model, in one of these equilibria the proportion of maroon balls must strictly exceed the proportion in the other equilibrium. Consider two paths, one that leads to the first equilibrium and one that leads to the second. Along the first path, an outcome in a given period may or may not have some effect on the equilibrium, the same is true along the second. But in some periods, these outcomes have to shift the probabilities of future outcomes. Theoretically, weight could be assigned to each outcome in terms of how much influence it has on future outcomes. These influences would have to exaggerate one outcome at the expense of another. In other words, the influences would have to reinforce one outcome and exclude the other, or put loosely the process would have to exhibit increasing returns. This is why people think path dependence requires increasing returns. The flaw in this logic is that the reinforcement

[9]The increase can be explained as follows: in each period there are fewer maroon balls than brown balls, so adding one maroon and one brown increases the relative proportion of maroon balls.

or exclusion need not be in the form of increasing returns. They could occur through complementarities as was true in the Balancing Polya Process.

That increasing returns implies path dependence seems obvious. Yet this inference is also logically incorrect. All of the outcomes could have increasing returns, but if one outcome has much stronger increasing returns than the others, it will always win. That is true of the brown balls in the Biased Polya Process. It was also true for gasoline powered automobiles.

There is a substantial amount of recent research showing how increasing returns and positive feedbacks can lead to path dependence. Much of this research shows that a particular institutional arrangement, e.g. rent seeking by oil-rich states (Karl 1997) or party patronage systems (Shefter 1977) comes into being and then because of increasing returns results in a path dependent outcome. The Karl example is instructive. It may be more accurately described as a case of negative externalities. The easy profits in the oil industry destroy the incentives for developing other industries. While it is true that oil begets oil, it is also true that oil precludes cars and semiconductors.

More generally, when someone finds evidence of increasing returns and positive feedbacks, whether it be with respect to technology or international institutions (Weber 1997) this evidence is not sufficient proof that multiple equilibria exist. It could well be that the system has a single equilibrium, that one configuration's positive feedbacks and increasing returns are so large as to swamp any others.

History offers an abundance of examples with increasing returns but no path dependence. Around the turn of the century, gas, steam, and electric powered cars competed in the marketplace. All three technologies exhibited the classic properties of increasing returns to scale technologies: falling production costs, fuel delivery systems, etc.. Of the three, gasoline powered cars had by far the largest increasing returns (you cannot carry electricity in a can), and they won the market. Rerun history ten times, a thousand times, even a million times, and the gasoline engine wins every time. Increasing returns? Yes. Path dependence? No. Similarly, the money raised in political campaigns, especially primary campaigns with a dominant incumbent, may also exhibit increasing returns but have a single equilibrium (at least qualitatively speaking). The incumbent may raise almost 100% of the money, but this does not deny that the other candidates' fund raising efforts exhibited increasing returns. They probably did, but their increases were just too small.

## 3.5 Strong Path Dependence

The Biased Polya process is path dependent. If in the first five periods one brown ball is selected, the number of balls in the urn depends upon the period in which it was selected. However, there are multiple paths that generate the same outcome probabilities. If in the fourth period, there are ten brown balls and three maroon balls, it could either be that a brown ball was selected only in period three, or that a brown ball was selected in periods one and two but not in period three.

It is therefore possible to construct an even stronger notion of path dependence, namely that *any two distinct paths lead to different outcome probabilities* I refer to this as *strong path dependence.*

*A process is* **strong path dependent** *if for any two distinct histories, the outcome function differs. A strong path dependent process can be written as follows:*

$$x_{t+1} = G_t(h_t)$$

Where $G_t(h_t) \neq G_t(\hat{h}_t)$ if $h_i \neq \hat{h}_i$ for some $i = 1$ to $t$.

Strong path dependence implies path dependence. Strong path dependence might also be called *order of the path* dependence. The next two examples show that it is possible to construct a strong path dependent process using the simple urn model, and that it is also possible (up to a set of paths of measure zero) to construct a strong path dependent process that is equilibrium dependent.

**Example 6: A Strong Path Dependent Process:** Initially $M = B = 1$. In period $t$, a ball is chosen and $2^{t-1}$ balls are added to the urn of the color of the chosen ball.

To see that this process is strong outcome path dependent, it helps to consider a specific path of outcomes, say $MBMMB$. After period 1, a maroon ball is added. After period 2, two brown balls are added. After periods 3 and 4, four and eight maroon balls are added, and after period 5, sixteen brown balls are added. Therefore, before the ball is chosen in period six there are fourteen maroon balls $(1+1+4+8)$ and there are nineteen brown balls $(1 + 2 + 16)$. It can be shown that $MBMMB$ is the unique history that generates fourteen maroon balls and nineteen brown balls.[10] Therefore, the process is strong path dependent. This process does not converge to any fixed probability selecting a maroon ball.

The next process, however, does converge at least up to a set of measure zero.

**Example 7: The Burden of History Process:** Initially $M = B = 1$. In period $t$, a ball is chosen and put back in the urn together with a ball of the same color. In addition, for each period $s < t$ , $2^{t-s} - 2^{t-s-1}$ balls are added to the urn of the color of the ball chosen in period $s$.

This process relies on the same basic construction as the previous example. After $T$ periods, there will be $2^{T-1}$ balls placed in the urn that match the ball selected in the first period, $2^{T-2}$ that match the ball selected in the second period and so on. In this process, the first ball selected always determines the color of approximately one half of the balls added to the urn , the second ball selected determines approximately one fourth

[10] Change $M$'s to 1's and $B$'s to 0's. This gives a binary sequence. The number of maroon balls in the urn equals the integer conversion of the binary sequence which is unique.

of the balls added to the urn and so on. Later periods matter exponentially less. The process can be shown to converge to a unique equilibrium distribution for any history up to a set of measure zero.[11]

This last process further clarifies the relationship between path dependence and increasing returns. The Polya Process notwithstanding, path dependence is loosely conceived as implying that the entire path matters. There is often an implicit assumption that the weight of early history, that early decisions, actions, and choices grow more and more powerful throughout time. In this last process, as time unfolds the past takes on more and more weight. In doing so it creates strong path dependent equilibria. These two loose conceptualizations: that of strong path dependence and that of an increasing weight of past history are linked. Even though increasing returns are neither necessary or sufficient for equilibrium dependence, some form of increasing returns, or increasing complementarities, can create path dependence and not just phat dependence. I return to this insight later in the paper, when I discuss the possibility of path dependence in the externality models.

### 3.5.1 Early and Recent Path Dependence

Some processes do not depend upon the entire path, but on the initial history or on the early path. In these cases the histories can be partitioned into a finite number of sets, so the processes can be thought of as state dependent. However, if the early path effects later outcomes, the these processes cannot satisfy the assumptions of the ergodic theorem. Typically, they violate ergodicity. Once as state has been attained it gets *locked in.*

*A process is* **initial outcome dependent** *all subsequent outcomes depend only on the first outcome. An initial outcome dependent process can be written as follows:*

$$x_t = G(h_2)$$

The next example describes an initial outcome dependent process in which a founder makes a random decision which subsequently charts the future course of events deterministically.

**Example 8: The Founder Process:** $M = B = 1$. If the ball chosen in period one is maroon, the maroon ball is put back in the urn and the brown ball is removed. Similarly, if the ball chosen is brown, the brown ball is put back in the urn and the maroon ball removed.

[11] It is possible in this process that two histories converge to the same probability distribution. For example, the history of one maroon ball and then all brown balls and the history of one brown ball and then all maroon balls both converge to equal probability of maroon and brown balls. But this is a zero probability event.

This process has only two paths. All future outcomes must be the same as the first outcome. This process is an example of strictly initial outcome dependence.

*A process is* **early path dependent** *if the outcome in any subsequent period depends only upon the history up to some period $T$. An early path dependent process can be written as follows:*

$$x_{t+1} = G_t(h_t) \text{ for } t \leq T, \text{ and } x_{t+1} = G(h_T) \text{ for } t > T$$

This phenomenon has been popularized in the information cascades literature (Bikhchandani, Hirshleifer and Welch 1992, 1998, Lee 1993) When enough consecutive people vote yes or buy a stock others follow like lemmings regardless of their information, so only the early part of the path matters. Example 9 provides a simplified version of a cascade.

**Example 9: A Cascade** Initially $M = B = 1$. Balls are selected and replaced in the urn until three consecutive balls of the same color are selected. When this occurs, the ball of the other color is removed from the urn.

In this process, market and bureaucratic outcomes are equally likely until three consecutive outcomes are identical at which point the process locks into that outcome. This captures the phenomenon that the early history matters exclusively. As previously mentioned, courts may set precedents that greatly influence future rulings. If so, the process of legal decisions may depend heavily upon the early path.

## 3.6 Recent Path Dependence

Finally, It is also possible that processes depend not on the early path but on the recent path. Retrospective voting processes would be an example of recent path dependence. The notion of recent path dependence runs counter to common conceptions of path dependence which emphasize early decisions. Many empirical investigations suggest the prevalence of recent path dependence. If the recent path matters, then the depth of a process can be measured as the number of periods back that influence the next state and outcome. This notion of depth is borrowed from the physical concept of thermodynamic depth.

In the interests of brevity, I construct examples of last outcome dependence and recent path dependence but only formally define recent path dependence.

*A process is* **recent path dependent** *if the outcome in any subsequent period depends only upon the outcomes and opportunities in the recent past. A recent path dependent process can be written as follows:*

$$x_{t+1} = G_t(h_t) \text{ for } t \leq T, \; x_{t+1} = G(h_t/h_{t-T}) \text{ for } t > T$$

The first example is the Unstable Government Process. There are two parties. Each party has a preferred institution $M$ or $B$. If the party's preferred institution is selected, the party stays in power. If not, the party falls out of power.

**Example 10: An Unstable Government Process:** There are two parties: $D$ and $R$. If $D$ is in power then $M = 1$ and $B = 2$. If $R$ is in power then $M = 2$, and $B = 1$. A randomly chosen party is in power in the first period.[12] Each period, a ball is chosen. In subsequent periods, the party in power equals $D$ if a $B$ was chosen and equals $R$ if an $M$ was chosen.

This process satisfies all of the conditions of the ergodic theorem, and in the unique equilibrium distribution, maroon and brown balls are equally likely to be selected. The second example I call a Forgetting Process. The process has finite memory; only the recent past matters.

**Example 11 A Forgetting Process** Initially $M = B = 1$. An additional ball of the same color as the selected ball is added for $K > 0$ periods and then removed from the urn.

The Forgetting Process does not only depend upon the last outcome, but on the recent past as well. In this way it differs from the Unstable Government Process.

## 3.7 Overlapping Definitions

The previous definitions do not create a classification since multiple definitions often apply to a single case. For example, a recent outcome path dependent process can also generate early path dependent equilibria. Consider an urn that begins with five maroon and five brown balls that are numbered from one to ten. In each new period, when a ball is drawn it is replaced and another ball of the same color is added. If in addition, the lowest numbered ball is removed from the urn, then this process generates both an early path dependent equilibrium as well as a recent path dependent equilibrium since at some point the urn contains only one color of ball.

In addition to structuring our thinking, the previous definitions allow for more precise empirical testing (Jackson 2004).[13] For example, provided data were available, phat dependence and path dependence demand different tests. In the former case, only the set of past outcomes would be used to predict the current outcome. In the latter, the order of those outcomes would also be included.

[12]Parties are each chosen with probability one half

[13]I thank John Jackson for suggesting this approach. See Jackson (2004) for how these various formalizations might be tested.

## 3.8 Cliff's Notes

Each of the eleven examples highlights a feature of dynamic processes related to path dependence. Those features are summarized in the table below.

| | | |
|---|---|---|
| 1 | Bernoulli | Independence |
| 2 | Polya | Phat Dependence: Multiple Equilibria |
| 3 | Balancing | Phat Dependent, Unique Equilibria |
| 4 | Balancing Polya | Phat Dependent Equilibria, Not Increasing Returns |
| 5 | Biased Polya | Increasing Returns, Unique Equilibria |
| 6 | Strong Path Dependent | Path Dependence, May Not Converge |
| 7 | Burden of History | Path Dependent, Converges |
| 8 | Founder | Initial Outcome Dependent |
| 9 | Cascade | Early Path Dependent |
| 10 | Unstable Government | Last Outcome Dependent |
| 11 | Forgetting | Recent Path Dependent |

# 4 Externalities, Changing Externalities, and Path Dependence

I now describe a second class of models based on externalities between actions and choices. The ball and urn models rely on minimal notation. They are visualized and understood with little effort. At the same time, they do not resemble much of the social, political, and economic world. They are severe abstractions. In an urn model, current and past outcomes (draws from the urn) increase and decrease the probabilities of future outcomes. That's okay. The urn models are meant to give intuition, and they are successful in providing it, but social science needs richer models of path dependence if it is to progress. We need models with agents who possess utility or objective functions, who make explicit choices. These models must include specific assumptions about the externalities between choices and actions that create the directional forces in the space of outcomes.

Models based upon externalities between choices and actions enable a deeper unpacking of the causes of path and phat dependence. These models are not new to the discipline, but it is fair to say that they are not common. The models I present here are similar to those of List (2003) though I abstract from the micro level detail about group preferences that is central to his analysis.

## 4.1 A Basic Model of Externalities

The model assumes a decision maker who takes sequential actions. These actions can be thought of as a infinite sequence of choices over project proposals such as whether to build a highway or bridge or over institutional decisions such as whether to create a post office or a national guard. The fact that the sequence is infinite implies that the decision maker cannot wait until she sees all possible projects prior to making a decision. For presentation purposes, hereafter I refer to each action as a decision. This decision could be either over a public project, a technology, or an institutional feature.

In each period, exactly one decision is considered. A decision is identified by the period in which it is first considered. An extended model in which multiple decisions arise in any period yields similar results.[14] I assume that an affirmative decision cannot be reversed, but a decision that is rejected can be accepted at any later period. This assumption captures the fact that once created, most public projects, technologies, or institutions cannot be undone without a cost. I am in effect, assuming that this cost is infinite. This is a strong form of irreversibility.

Without some type of irreversibility, history dependence would be less likely to occur because anything done could be undone – at any point in time, a decision maker could undo the past and implement the optimal set of choices. In the current model,

[14] Multiple periods in the current model can be condensed to a single period in a more general model.

the assumption of infinite costs can be relaxed without effecting the main result that history dependence is not implied by positive externalities or increasing returns but the mathematics would be more cumbersome.

The value of the decision proposed in period $t$ equals its *isolated value*, $v_t$, which may be either positive or negative plus the value of any externalities it generates. I assume that this value and the value of the externalities are known. That of course is problematic (Jervis 1997). In the complex real world, the values of decisions are rarely known with certainty. I assume here that all externalities are among pairs of decisions, but the results still hold if I admit higher order externalities. I let $e_{st}$ denote the externality between the decision in period $t$ and the decision in period $s$, where $s < t$. For the moment, I consider the $e_{st}$'s to be positive, negative, or of zero value. It is important that externalities are only realized among decisions that are approved. Suppose that the beginning of the fourth period that the decisions considered in period one and period three have been approved but that the decision considered in period two was rejected. Even though the decision considered in period three generates externalities $e_{13}$ and $e_{23}$, only the former is realized if the decision is accepted. As a result, the value of the set of approved decisions after period three equals $v_1 + v_3 + e_{13}$, and the marginal value of adding the decision considered in period four equals its isolated value, $v_4$, plus the the sum of the values of the externalities it creates with decisions one and three, $e_{14} + e_{34}$.

The first conclusion to draw from this model is that without externalities there is no potential for history dependence. Each decision can be made in isolation. Previous decisions have no effect on current decisions. With externalities, earlier decisions can constrain or influence later decisions and therefore the sequential decision making process can be history dependent. Consider the following model specific definition of a history dependent decision rule. For a given finite set of decisions, an *ordering* of those decisions is a sequence and a *permutation* of that ordering is a rearrangement of that sequence.

*A decision rule is* **history dependent** *if there exists a reordering of some finite set of decisions that yields a different set of approved decisions.*

This is an analog of my previous definition of history dependence redefined within this model. The outcome $x_t$ can be thought of as the set of decisions. The definition requires that the history of decision proposals influences the set of decisions approved. There will always exist rules that are not history dependent, such as a rule that accepts or rejects all decisions. But neither of those rules would necessarily be a good rule to follow. The relevant question is whether there exists a decision rule that makes optimal choices which is not history dependent. As the sequence of decisions is infinite, this necessitates a clarification of what is optimal. I could define optimality with respect to the present discounted value of decisions. I could also define optimality with respect to each period. The main results of this section hold given either convention. I choose to adopt the latter convention of optimality because it makes the analysis more transparent.

Define the *acceptance set in period* $t$ as the set of decisions approved in the first $t$ periods. I can then define optimality as follows:

*A decision making rule is* **optimal** *in period $t$ if the acceptance set in period $t$ obtains the highest possible value for all subsets of the first $t$ decisions.*

This formulation has a political interpretation. If an election were held after period $t$, the leader would want to have approved the best set of decisions among the $2^t$ possible subsets of $t$ decisions. This construction implies that a decision rule that is optimal up to period $t$ need not be optimal up to period $t+1$. If a decision arises in period seven that creates negative externalities with decision three, then the optimal set of decisions at the end of period six may include decision three and not decision seven. But if decision seven has a high isolated value, then the optimal set of decisions at the end of period seven may include decision seven and exclude decision three. Given that acceptances are irreversible, once decision three is approved, it cannot be undone. Therefore, any rule that is optimal in period six cannot be optimal in period seven.

To avoid the messiness that arises with multiple optimal acceptance sets, I assume that for any $t$ there is a unique optimal acceptance set. Without this restriction, it is possible to construct examples in which the decision rule is history dependent, but the paths chosen have the same value and therefore history is irrelevant.[15] I can now state the following observation.

**Observation 5** *A decision making rule that is optimal in every period cannot be history dependent*

The proof is by contradiction. Suppose that a decision making rule is history dependent. It follows that it cannot be optimal for all $t$. Consider an ordering of decisions and any permutation of that ordering that generate distinct acceptance sets in some period $\hat{t}$. In period $\hat{t}$ only one of these acceptance sets can be optimal, a contradiction.

Showing that optimality and history dependence cannot coexist does not settle matters. It remains to show that there exists an optimal decision rule in nontrivial contexts, otherwise the observation is vacuous, but that is easily accomplished. Assuming only positive externalities, the optimal decision rule in period $t$ is as follows: (i) accept the decision in period $t$ if it has a positive isolated value (ii) accept the decision in period $t$ if its isolated value plus its externalities with all approved decisions is positive and (iii) after the decision to accept or reject the decision in period $t$ has been made using criteria (i) and (ii) accept any subset of previously rejected decisions if when approved collectively, they increase the value function.

To see how to apply this rule, suppose that each of the three decisions has an isolated value of negative 5 but that each generates a positive externality of 6 with every other

[15]This assumption can be relaxed in several ways, but I want to make the intuition as clean as possible.

decision. Decision one considered alone would be rejected using this rule. Decision two considered alone and as a pair with decision one would also be rejected. Decision three would be rejected when considered alone but when considered with both decisions one and two, the three of them would be approved because the create a total value of three (6+6+6 -5-5-5). The intuition behind the optimality of this rule is straightforward. Since all externalities are positive, any decision or subset of decisions that contributes positively to the total value in period $t$ will continue to do so as more decisions are added. This verbal argument can be formalized.

**Observation 6** *With only positive externalities between decisions, there exists a decision making rule that is optimal up to period t for all t.*

For a proof of this observation see Page (1997). A corollary to this observation is that increasing returns need not imply a path dependent decision rule.

**Corollary 1** *With only positive externalities between decisions then there exists an optimal decision making rule, which is therefore not history dependent.*

The next relate this result to environments with increasing returns. I first define increasing returns using the externalities framework I have constructed. To accomplish this, I create types. These types represent the characteristics of decisions or institutions upon which there may be increasing returns. I first assume that each decision belongs to one of a finite number of types: $\{A, B, C...L\}$. I then say that a collection of decisions exhibits *exclusively increasing returns* if decisions belonging to the same type create positive externalities with one another but create no other externalities. Formally, this means that $e_{st}$ is positive if the decisions considered in periods $s$ and $t$ belong to the same type and is zero otherwise. The next two observations follow from the observation above and its corollary.

**Observation 7** *If each decision is a single type and the decisions exhibit exclusively increasing returns then there exists an optimal decision rule which is therefore not history dependent.*

Second, assume that decisions can have multiple types, but that again there are exclusively increasing returns.

**Observation 8** *If decisions have multiple types and if the decisions exhibit exclusively increasing returns then there exists an optimal decision rule which is therefore not history dependent.*

These two observations prove that increasing returns alone does not create history dependence. However, in the sequential decision making context, the presence of negative externalities makes path dependence difficult to avoid. The reason for this is obvious. Negative externalities imply that once a decision is made, future decisions are constrained in a way that compromises optimality. This can be stated formally.

**Observation 9** *If there does not exist an optimal decision making rule for some subset of decisions for all periods up to period $t$, then there exists at least one negative externality between decisions.*

For a proof of this observation see the appendix.

## 4.2 QWERTY

With these observations in mind, I return to the example of the QWERTY typewriter keyboard. In the great sweep of history, keyboard configurations are significantly less than a minor footnote, the QWERTY example is by far the most well known example of path dependence and it is one reason people confuse path dependence and increasing returns. What follows is a brief recapitulation of the QWERTY history and an explanation that negative externalities are the cause of the path dependence.

Though there are many possible keyboard arrangements, for a variety of reasons, initially only QWERTY typewriters were offered for sale. People then could only buy typewriters with QWERTY keyboards.[16] As typing on a keyboard with a different key configuration requires learning to type anew, as more people bought QWERTY keyboards, QWERTY keyboards became locked-in. Thus, QWERTY keyboards exhibit increasing returns and they created a path dependent process.

This simple version contains a subtle but important blurring of the causes of path dependence. While it is true that the QWERTY typewriter keyboard exhibits increasing returns, these are increasing *relative* returns not increasing absolute returns. The QWERTY typewriter becomes relatively more valuable than other typewriters because it creates both positive and negative externalities. The positive externality is between QWERTY keyboards. The more people who could type on QWERTY keyboards, the more valuable a QWERTY keyboard becomes as other QWERTY typists can also use them. But, the sale of more QWERTY typewriters also creates two negative externalities. Once a person has one typewriter, she derives little benefit from having another one, especially one with a different keyboard configuration. This is an *intra personal* negative externality. Second, as more people buy QWERTY keyboards, the value to a person of learning on another keyboard decreases because the prevalence of QWERTY keyboards. This is an *interpersonal* negative externality.

Therefore, the increasing relative returns to QWERTY are due to both *positive externalities* for QWERTY typewriters and *negative externalities* imposed on other typewriters. The same logic applies in the case of VHS and BETA technologies. As more people bought VHS machines, this created positive externalities for future VHS purchasers (more available titles) and negative externalities for future BETA purchasers (relatively fewer available titles).

[16] My favorite explanation for the QWERTY configuration is that it allowed people to type the word *typewriter* without moving their fingers from the top row of keys.

This decomposition of increasing relative returns into positive and negative externalities enables a recasting of the QWERTY example in which the negative externalities can be seen as the primary cause of path dependence. To show this, I construct a simple model with two keyboard configurations: the QWERTY keyboard, $Q$, and the ZRJSOC keyboard, $Z$. Let $N_Q$ and $N_Z$ denote the number of people who have bought each type and let $x_i \in \{0, Q, Z, B\}$, denote whether a person $i$ has purchased no typewriter, a QWERTY typewriter, a ZRJSOC typewriter, or both typewriters. In this simple model each person has a value for each keyboard. The values depend upon the number of people who have bought each type as well as whether or not the person has bought a typewriter. Thus, person i's values for the typewriters with QWERTY and ZRJSOC keyboards can be written as $V_{iQ}(N_Q, N_Z, x_i)$ and $V_{iZ}(N_Q, N_Z, x_i)$, respectively. There are also prices $p_Q$ and $p_Z$ for the QWERTY and ZRJSOC typewriters.

I assume that there are a large number of periods and that each period, a salesperson shows up at a random person's door offers that person a typewriter with either a QWERTY or a ZRJSOC typewriter. This choice over keyboards is also assumed to be random. I first assume only positive externalities. If the first person buys a QWERTY typewriter, this increases the value of QWERTY typewriters for other people. This first purchase may even cause an initial cascade of QWERTY purchases. However, it has no effect on peoples' values for or the price of the ZRJSOC typewriters. Therefore, at some point, those people whose values for the ZRJSOC typewriter exceed the price will purchase ZRJSOC typewriters as well. This in turn induces others to buy them. In the long run, ZRJSOC purchases are unaffected by the path of purchases. So, although, the outcome in any period may depend upon the path, but the long run distribution over outcomes does not. The process is not path dependent.

To see this in an example, assume that there are five people numbered from one to five and indexed by $i$. Let $V_{ix} = 2i + 2N_x$ for $x = Q, Z$. Assume that the price of any typewriter equals \$5. Initially, persons 3, 4, and 5 would buy either a QWERTY or a ZRJSOC typewriter, but the other two people would buy neither. Suppose that person 5 buys a QWERTY typewriter. Now, person 2 would also be willing to buy a QWERTY typewriter, even though person 2 would not be willing to buy a ZRJSOC. If any two people buy QWERTYs, then person 1 will also want to buy a QWERTY, thus eventually, everyone owns a QWERTY. However, by the exact same logic, eventually everyone owns a ZRJSOC as well.[17]

To see this logic in another way, suppose that the two products are VHS machines and QWERTY typewriters. Each creates positive externalities within its own type but no externalities with the other type: buying a QWERTY typewriter creates no negative externality for a VHS machine, and buying a VHS machine creates no negative

[17]If each person is only offered one typewriter to buy, then path dependence would occur. If the first person buys a QWERTY, then this raises the value of QWERTYs relative to the ZRJSOC typewriters, and the result would be path dependent. However, the assumption of a single purchasing opportunity is equivalent to a negative externality.

externality for a QWERTY typewriter. Therefore, the number of QWERTY typewriters and VHS machines sold does not depend at all on the path of purchases.

Assume that there are only interpersonal negative externalities. The case for intra personal works similarly. Let $V_{iQ} = 2i - 2N_Z$ and $V_{iZ} = 2i - 2N_Q$. Now, when a person buys a QWERTY typewriter, she lowers the value of ZRJSOC typewriters for others. Consider two scenarios. In the first scenario, person 5 and person 4 both buy QWERTY typewriters. Person 3's value for a ZRJSOC typewriter falls from \$6 to \$2, so person three will now only buy a QWERTY typewriter. Suppose person 3 buys a QWERTY typewriter. Person 5's value for a ZRJSOC typewriter now falls to \$4, and person 4's value falls to \$2, so no more sales of ZRJSOC typewriters occur. In the second scenario, person 5 buys a QWERTY typewriter, and person 4 buys a ZRJSOC typewriter. Now, person 3 buys neither a QWERTY or a ZRJSOC typewriter owing to the negative externalities. His value is now \$4 for each type. Person 5 however also purchases a ZRJSOC typewriter. Person 4 does not purchase a QWERTY typewriter, because two people own ZRJSOC typewriters reducing person 4's value for a QWERTY to\$4. Thus, these two different paths lead to different equilibria.

Thus, inter personal positive externalities do not create path dependence, but that interpersonal negative externalities do. None of this implies that positive externalities do not contribute to path dependence. They do. Positive interpersonal externalities reinforce the path dependence. If I include both types of externalities $V_{iQ} = 2i + 2N_Q - 2N_Z$ and $V_{iZ} = 2i + 2N_Z - 2N_Q$. Now, if persons 3 , 4, and 5 buy QWERTY typewriters, everyone buys a QWERTY and no one buys a ZRJSOC. Thus, if both positive and negative externalities exist, then path dependence will be powerful and ubiquitous. The positive externalities create a bias in favor of some decisions and the negative externalities create a bias against others. Both forces operate in the QWERTY example, which is why there is such exaggerated path dependence. Yet, to refer to this as a case of increasing returns causing path dependence is misleading given that the negative externalities are the true cause and the positive externalities only exaggerate the phenomenon.

## 4.3 Negative Externalities and Constraints

If negative externalities create history dependence, where do they come from? One answer is constraints. These constraints can be with respect to domain, time, money, or space. Regardless, they create negative externalities. Any large public decision, be it a prison or university takes up space and requires money. Both create negative externalities with future public projects. Obviously, the more money and space project demands, the greater its impact on the path. Small projects are less likely to influence the path of history than are large projects. That is not to say that smaller decisions cannot accumulate over time and restrict history to certain paths, but that any big project crowds out other projects.

Power also creates constraints. A president has only so much political capital. He (or she) must decide where to spend it. Whatever choices are made may exclude other opportunities. Again, this can result in path or phat dependence. Finally, cognition imposes constraints. Our heads only have room for so much. We borrow behaviors from one setting and apply them to another. This consistency is part of culture and it can create history dependence in ways that simple externalities cannot a point I take up next.

## 4.4 Path and Phat Dependence and Externalities

A further implication of the externality framework as formulated is that to the extent that history dependence exists, it is phat dependence not path dependence. This is because a set of decisions defines a set of externalities and those externalities are independent of the order in which the decisions were approved. The order of decisions does not effect the size or scope of the externalities. To make this observation, I distinguish between the order in which decisions were approved.

*A decision making rule is* **independent** *if it does not depend on the set of approved decisions. It is* phat dependent *if it depends upon the set of approved decisions, and it is* path dependent *if it depends upon the order in which decisions were approved.*

Given these definitions, I can make the following observation.

**Observation 10** *If the magnitude or number of externalities associated with a decision does not change over time, then a decision rule need only be phat dependence. It cannot benefit by considering the order of the path.*

If then we want to make the case for path dependence in a world with externalities, we need those externalities to accumulate or abate over time. Either would result in the path of decisions mattering. This still does not imply as much path dependence as might be though. For example, there is evidence of increasing attachment of how political and economic institutions complement existing institutions (Ikenberry 2001). Thus, the longer an institution has been around, the more time these complementary institutions have had to spring up. Yet, this does not necessarily imply path dependences as the set of institutions differs depending upon the order. It is more likely *phat dependence.*

This does not imply that there cannot exist path dependence. In the case of projects, institutions, and technologies, the build up of cognitive and behavioral repertoires are a likely cause of path rather than phat dependence. Following the lead of North (1991) and Grief (1994), the possibility exists not only for richer historical analysis Pierson (2004) Chong (2000) but also for theoretical models of how environments shape behavior and beliefs (Bednar and Page 2004). These histories and models that make explicit linkages between past institutional choices to current beliefs and behavior.

## 5 Concluding Thoughts

In this essay, I have provided definitions of various types of path dependence. I have distinguished between processes that are outcome path dependent, in which history only matters for the current outcome, and those that generate path dependent equilibria, in which history determines the course of future events, and I have distinguished between path dependence, in which the order of the history matters, phat dependence, in which only the set of historical events matters, and state dependence, in which history can be classified into finitely many bins. I have also shown a lack of equivalence between path dependence and increasing likelihood of outcomes. This focus on increasing returns and positive externalities is misleading. Negative externalities are more appropriately seen as the driving force behind path dependence.

If negative externalities are ubiquitous and are a cause of path dependence, why has there been a fascination with positive externalities as a cause of path dependence and a neglect of negative externalities? First off, negative externalities are neither new nor sexy. Social scientists have long been aware that budget, time, and spatial constraints require choosing between options and that these choices constrain future choices. Second, stories about QWERTY keyboards, Microsoft software, and tropical economies are compelling and accurate examples of extreme path dependence. Third, as my reconsideration of the QWERTY example shows, positive externalities exaggerate the degree of path dependence. Given the obvious importance of path dependence, the focus on extreme cases is warranted.

Of course, the existence of negative externalities need not imply path dependence either. They are necessary but not sufficient for path dependence. If the decision maker has perfect foresight, then she may be able to make an optimal set of decisions, though it may not appear optimal at a particular point in time. For path dependence and rational choice to coexist the externalities that make the path matter must be unanticipated or unanticipatable because path dependence hinges upon a lack of perfect foresight. In some instances, the time scale may be sufficiently long to make the perfect foresight assumption unlikely. The standard QWERTY example of coordination forcing a system down a particular path requires that the people making decisions cannot see down the entire path or do not have any incentive to do so. It is not unreasonable to assume that typewriter designers could not anticipate the electronic keyboard or the social benefits from a design that optimized speed (Liebowitz and Margolis 1990).

The intuition derived in the previous section that path dependence and optimality are at odds with one another suggests List's (2003) complementary result that a lack of internal consistency also can lead to path dependence. List's results are important in light of the fact that many large decisions are made by collectives and not by individuals. While an individual may be consistent, no such claims can be made for groups, especially groups with changing compositions. Even forward looking political bodies may make path dependent decisions in sequential contexts provided the decision problems are difficult. So, those contributing to growing literature on how institutions may generate path dependent outcomes should do careful accounting of externalities and the complexities they generate.

In conclusion, this essay was intended to be introductory not definitive. We have not yet begun to write models and frameworks for path dependence. Someday the ball and urn models of path dependence may look as primitive as the hydraulic models of the economy that people used to build. That is not to say that we should ignore them or relegate them to the undergraduate classroom, but we should see the ball and urn models for what they are, simple models that clarify our thinking. It is to say that we have an opportunity to construct richer, more realistic models and that these models can help us better understand the many ways that histories unfold.

# Appendix

Proof of observation that negative externalities are necessary for path dependent decision rules.

By assumption there exists a period $s$ such that the optimal acceptance set at time $s$, $X_s^*$, is not a subset of the optimal acceptance set at time $s+1$, $X_{s+1}^*$. Let $V(X)$ denote the value of the subset of decisions $X$. Given that there is a unique optimal acceptance set, it follows that

$$V(X_s^*) - V(X_s \cap X_{s+1}^*) > 0$$

Suppose that all externalities are positive. It follows that for any subset of decisions denoted by $A$

$$V(X_s^* \cup A) - V((X_s^* \cap X_{s+1}^*) \cup A) > 0$$

This must be true for $A$ equals $X_{s+1}^*$ Therefore,

$$V(X_s \cup X_{s+1}^*) - V(X_{s+1}^*) > 0$$

which contradicts the optimality of $X_{s+1}^*$.

## References

Arthur, Brian (1994) *Increasing Returns and Path Dependence in the Economy.* The University of Michigan Press, Ann Arbor.

Bednar, Jenna (2005) *The Robust Federation* book manuscript.

Bednar, Jenna and Scott Page (2004) "Can Game(s) Theory Explain Culture? The Emergence of Cultural Behavior Within Multiple Games " Santa Fe Institute Working Paper #04-12-039.

Bikhchandani, Sunil, David Hirshleifer and Ian Welch (1992) "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy* 100:5, 992-1026

Bikhchandani, Sunil, David Hirshleifer and Ian Welch (1998) "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades" *The Journal of Economic Perspectives*, Vol. 12, No. 3, pp. 151-170.

Chong, Dennis (2000)*Rational Lives: Norms and Values in Politics and Society.* Chicago: University of Chicago Press.

Cowan, Robin, Philip Gunby (1996) "Sprayed to Death: Path Dependence, Lock-in and Pest Control Strategies." *The Economic Journal*, Vol. 106, No. 436, pp. 521-542.

Crouch, Colin, and Henry Farrell (2004) "Breaking the Path of Institutional Development? Alternatives to the New Determinism"*Rationality and Society* vol 16(1) pp 5 -43.

David, Paul (1985) "Clio and the Economics of QWERTY" *The American Economic Review, Vol. 75, No. 2, Papers and Proceedings of the Ninety-Seventh Annual Meeting of the American Economic Association.* pp. 332-337.

Easterly, W. (2001). *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics.* Cambridge, MA: MIT Press.

Ferejohn, John. (1991) "Rationality and Interpretation: Parliamentary Elections in Early Stuart England," in K.R. Monroe, ed., *The Economic Approach to Politics: A Critical Reassessment of the Theory of Rational Action.* New York: HarperCollins.

Gaddis, John L. (2002) *The Landscape of History: How Historians Map the Past* Oxford University Press.

Gerschenkron, Alexander. (1952). "Economic Backwardness in Historical Perspective." in *The Sociology of Economic Life*, edited by Mark Granovetter and Richard Swedberg. Boulder, CO: Westview Press.

Gilovich, T, Vallone, R., and Tversky, A. (1985)." The hot hand in basketball: On the misperception of random sequences." *Cognitive Psychology*, 17, 295-314.

Grief, Avner (1994) "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *The Journal of Political Economy* Vol. 102, No. 5. pp. 912-950.

Hacker, Jacob (2002) *The Divided Welfare State: The Battle over Public and Private*

*Social Benefits in the United States* Cambridge University Press.

Ikenberry, G. J. (2001). *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars*. Princeton, NJ: Princeton University Press.

Jackson, John (2004) Personal Communication.

Jervis, R. (1997). —em System effects: Complexity in Political and Social Life. Princeton, NJ: Princeton University Press.

Karl, Terry (1997) *The Paradox of Plenty: Oil Booms and Petro-States* Berkeley: University of California Press.

Kollman, Ken (2003) "The Rotating Presidency of the European Council as a Search for Good Policies" *European Union Politics* 4 (Mar): 51-71.

Lee In Ho (1993). "On the Convergence of Informational Cascades." *Journal of Economic Theory* 61:2, 395-411

Liebowitz, S. J. and Stephen E. Margolis, (1990) "The Fable of the Keys," *Journal of Law and Economics*, 1990, 33: 1-27.

Liebowitz, Stan J. and Stephen E. Margolis (2002) *The Economics of Qwerty: Papers by Stan Liebowitz and Stephen Margolis*, edited by Peter Lewin, MacMillan NYU Press.

List, Christian (2003) "A Model of Path Dependence in Decisions Over Multiple Propositions" working paper London School of Economics.

North, Douglas C. (1990) *Institutions, Institutional Change, and Economic Performance* Cambridge University Press. Cambridge.

North, Douglas C. (1991) *The Journal of Economic Perspective*, Vol. 5, No. 1. pp. 97-112.

Page, Scott (1997) "An Appending Efficient Algorithm for Allocating Public Projects with Complementarities," *Journal of Public Economics* **64**:3 pp 291–322.

Page, Scott (1998) "On the Emergence of Cities" *Journal of Urban Economics* 45: 184–208,

Page, Scott (2005) "The Rule of Six" mimeo Univerisity of Michigan.

Pierson, Paul (2000) "Path Dependence, Increasing Returns, and the Study of Politics," American Political Science Review, Vol. 94, No. 2, June, pp. 251-67.

Pierson, Paul (2004) *Politics in Time: History, Institutions, and Social Analysis* Princeton University Press, Princeton, New Jersey.

Shefter, Martin (1977) "Party Patronage: Germany, England, and Italy." *Politics and Society* Vol 7, pp 403-452.

Van Evera, S. (1998). "Offense, defense, and the causes of war. *International Security*, 22, 5-43.

Weber, S. (1997). "Institutions and change." In M. W. Doyle and G. J. Ikenberry, (Eds). *New Thinking in International Relations Theory.* Boulder, CO: Westview Press.