LEFT, RIGHT, LEFT: INCOME DYNAMICS AND THE EVOLUTION OF POLITICAL PREFERENCES OF FORWARD-LOOKING BAYESIAN VOTERS

John Morrow†  
University of Wisconsin, Madison

Michael Carter††  
University of California, Davis

1 December 2009

Abstract. The political left turn, which lagged Latin America’s transition to liberalized market economies by a decade, challenges conventional economic explanations of voting behavior. This paper provides a theoretical framework to help understand these complex political-economic dynamics. To do this, we first build on forward-looking voter models and analyze political preferences under general families of income transition functions. We show that non-concave functions, which offer no prospect of upward mobility for segments of the population, may result in stronger support for redistributive policies than might otherwise be anticipated. Interestingly, numerical analysis of the model based on estimated transition functions, suggests much stronger support for redistribution than actually materialized over the first decade of economic liberalization. We thus eschew the assumption that voters had full information on their new economic reality, and model voters as Bayesians learners. We show that starting from a prior that was consistent with the so-called Washington Consensus vision of liberalization, voters would be expected to exhibit the sort of political dynamics observed in most of Latin America over the last two decades.

Keywords: income dynamics, redistributive politics, polarization, Bayesian learning, Latin America.

JEL Codes: D31, D72, D83, P16.

† Department of Applied and Agricultural Economics, morrow1@wisc.edu
†† Department of Agricultural and Resource Economics, mrcarter@ucdavis.edu

Acknowledgments. We thank Jorge Aguero, Yasushi Asako, Brad Barham, Swati Dhingra, Scott Gehlbach, Aashish Mehta, Hiro Miyamoto and Shiv Saini for insightful comments as well as seminar participants at UW-Madison Development and Political Economy Seminars and the Midwest International Economic Development Conference. CEDLAS provided excellent data through SEDLAC. The usual disclaimer applies.
1. Introduction

Most Latin American countries had transitioned to market economies by the early 1990’s. The largely center right political leadership that instituted these transitions continued to win national elections and persisted in power throughout the 1990s and into the early 2000’s. Since that time, electoral politics have turned sharply left. The recent suite of presidential elections have seen left-leaning candidates defeat more conservative opponents in Brazil, Bolivia, Chile, Ecuador, El Salvador, Nicaragua, Paraguay, Peru and Venezuela. Not only have these elections ushered in a political shift, they have in many instances been hotly contested by candidates offering fundamentally different economic visions. The goal of this paper is to provide a theoretical framework to help us understand the economic forces that underlie these complex political dynamics.

The influential body of political economy literature that focuses on economic inequality as a force that determines both political institutions and voting patterns would seem to offer a window into these political patterns (Acemoglu and Robinson, 2006; Boix, 2003). However, the fact that inequality measures tend to be remarkably stable over time makes it unlikely that inequality can explain the right-left voting dynamics of Latin America. A recent paper by Robert Kaufman confirms the inconvenient empirical fact that existing measures of economic inequality do a very poor job of explaining both political institutions and voting patterns in Latin America (Kaufman, 2007).

Although we could abandon the search for economic explanations of contemporary voting patterns and appeal to other factors to explain changing voting patterns, we instead take our cue from Benabou and Ok (2001) who model voters as forward-looking agents who (1) formulate their political preferences based on how redistributive schemes will influence their future stream of well-being; and (2) have full information about their economy’s income distribution dynamics. From this perspective, voters should be driven by income dynamics, not by the current level of income inequality or other features of the contemporaneous income distribution.

At first glance, this analytical direction might seem odd given that Benabou and Ok provide theoretical foundations for surprisingly conservative voting behavior in the face of high inequality,

1While the contemporary Latin American left cannot be defined by a shared economic model, this new left does share a largely populist impulse and desire to shift resources and opportunity to those at the bottom of the income distribution.
2For a type of hybrid approach, see for instance Chapter 10 of Roemer (2001).
3In a recent paper, Gary Fields makes this point even more strongly as he shows how economic inequality can actually increase during the early stages of a period of upward mobility that would surely dampen political preferences for redistribution (Fields, 2007).
not for the surprising shift towards redistributive preferences observed in Latin America. Benabou and Ok specifically show that concave income distribution dynamics that offer the prospect of upward mobility (or, POUM) can account for anti-redistribution conservatism. Under POUM, forward-looking voters who would benefit from redistribution in the short run do not benefit in the long run and therefore vote against long term redistributive policies. A first contribution of this paper is to generalize the class of income transition functions considered by these authors. We show that non-concave income transition functions of the sort suggested by poverty trap theory, which offer no prospect of upward mobility (or, NoPOUM), can result in a surprisingly and increasingly progressive electorate.

In an effort to corroborate this theoretical intuition, we estimate income dynamics in Latin American countries. These estimates indeed identify the sort of NoPOUM dynamics that would be expected to generate an increasingly pro-redistribution electorate. Surprisingly, applying these dynamics to our full information, forward-looking voting model indicates that the demand for redistribution should have been stronger and should have occurred well in advance of the recent suite of Latin American presidential elections.

A second contribution of this paper is to explore this second puzzle. We argue it is the assumption that voters have full information about their economy’s income distribution dynamics that is most problematic, especially in transition economies where the electorate had little prior experience of a liberalized market economy. In such circumstances, voters have little choice but to fall back on ideological priors about how such an economy might work. In Latin America, the shift to the liberal economic model was sold on the grounds that it would boost incomes and well-being in the lower half of the income distribution. In an effort to better capture political dynamics, we model voters as Bayesian learners who begin with this “POUM prior,” and then experientially update their priors based on their own stochastic income realizations. We show that this model of forward-looking, Bayesian voters offers an empirically tenable explanation of the recent right to left political evolution in Latin America.

The remainder of this paper is organized as follows. Section 2 develops a basic framework for individual and aggregate income dynamics in the presence of transient shocks, and models political support for redistributive policies by both myopic and forward looking voters who enjoy full information about the income dynamic process. Section 3 then introduces a general class of income transition functions. Section 4 discusses the implications of our analysis for the political economy of redistribution in Latin America. In this paper we concentrate on their application to the redistributive preferences of Latin American voters. 

4See for example (Przeworski [1991]. Our approach to modeling ideology as an idiosyncratic evolving process echoes (Bates et al. [1998] as a means to complement cultural and ideological political theories with rational choice.
distribution dynamics that include both concave (POUM) and poverty trap (NoPOUM) dynamics as special cases, and derives results on the political preferences of fully-informed forward looking voters. The analysis of Section 3 is applied to Latin American income dynamics in Section 4. Section 5 then relaxes the full information assumption, and explores political dynamics as voters learn about the true income distribution dynamics that characterize their economy. Section 6 shows this model of forward-looking Bayesian voters who confront a NoPOUM world can give rise to the political polarization and sudden political shifts that have been observed in 21st century Latin America.

2. Forward-looking Voters and the Demand for Redistribution

This section lays out machinery needed to discuss changing patterns in majoritarian voting when the electorate can choose among income redistribution schemes. The setting is a continuum of voters whose income evolves over time and fluctuates with idiosyncratic shocks each period. We consider the fraction of the voter population who rationally prefer income redistribution, which we term the demand for redistribution.

Each individual voter considers their individual income path, aggregate income of the economy and the longevity of a proposed redistributive scheme. Changes in the economy over time thereby induce changes in voting patterns, and support for a given policy is dependent on expected economic conditions while in effect. For ease of comparison to Benabou and Ok (2001), we preserve their notation where possible.

2.1. Stochastic Income Transitions. Individuals are indexed by \( i \in [0, 1] \) and have initial incomes \( X \equiv [0, \bar{y}] \). The initial distribution of income is given by a absolutely continuous cdf \( F_0 \) with mean \( \mu_0 \). Individual \( i \)'s income at time \( t \) is denoted by \( y_{it} \) and evolves according to the following relationship:

\[
y_{it+1} = I(d_t(y_{it}), \epsilon_{it+1})
\]

(2.1)

where \( I \) is a continuous, increasing function relating current income to last period’s income, and \( d_t \) is a deterministic function around which income fluctuates each period due to the current period shock \( \epsilon_{it+1} \). Note that under this specification, the impacts of \( \epsilon_{it+1} \) are transitory, affecting contemporaneous income, but not future income. This simplifying assumption makes the evolution of an individuals income independent of his or her past history of shocks.
We further assume that the deterministic function is simply the period 0 expectation of time $t$ income ($d_t(y_{i0}) \equiv E_0[y_{it}]$) and that the shocks $\epsilon_{it}$ are identically distributed. Equation (2.1) then reduces to:

\[ y_{it+1} = I(E_0[y_{it}], \epsilon_{it+1}) \]  

(2.2)

Defining $f(y) \equiv E[I(y, \epsilon)]$ as the expected income next period given $y$, and taking expectations of each side of Equation (2.2) yields

\[ E_0[y_{it}] = E_0[I(E_0[y_{it-1}], \epsilon_{it})] = f(E_0[y_{it-1}]) \]  

(2.3)

Equation (2.3) shows $f(y)$ maps expected income this period to expected income next period, and recursion shows $E_0[y_{it}] = f^{(t)}(E_0[y_{i0}]) = f^{(t)}(y_{i0})$. For this reason we refer to $f$ as the income transition function. Substituting into (2.2) yields\footnote{The assumptions leading to Equation (2.4) are stated as Assumption 1 which we maintain throughout:}

\[ y_{it+1} = I(f^{(t)}(y_{i0}), \epsilon_{it+1}) \]  

(2.4)

Equation (2.4) shows that conditional on initial income, current income fluctuates around expected income $f^{(t)}(y_{i0}) = E_0[y_{it}]$ with the same distribution.

2.2. “One-shot”/Myopic Demand for Redistribution. We now consider the political preferences of pocketbook voters whose incomes evolve according to a known income transition function $f$ of the type just outlined. Pocketbook voters choose policies which maximize their income, and for simplicity we assume they are risk neutral. Following Benabou and Ok (2001), we define a simple class of redistribution schemes composed of a flat tax $\tau$ and a lump sum transfer to all voters. Such redistribution schemes are denoted $r_\tau(y|F)$ where if $r_\tau$ is enacted in period $t$, each voter $i$ receives income $r_\tau(y_{it}|F_t)$ where $r_\tau$ is explicitly given by Equation (2.5).

\[ r_\tau(y_{it}|F_t) \equiv (1 - \tau) \cdot y_{it} + \tau \cdot (1 - D) \mu_t \]  

(2.5)

Here $\mu_t$ denotes the mean income of the population at time $t$ and $D$ denotes any dead weight loss under the redistributive scheme. Of primary interest are the “laissez faire” scheme $r_0$ and the
“complete equalization” scheme \( r_1 \). In any case, it is clear the poor will be more likely to prefer \( r_1 \) to \( r_0 \).

Election timing plays an important role in the demand for redistribution. Consider a majoritarian vote taken between \( r_1 \) and \( r_0 \) in period \( t \). Also suppose the vote is \textit{ex post} in the sense that the vote is taken after the realization of the shocks \( \epsilon_i \). Then voter \( i \) prefers \( r_1 \) to \( r_0 \) exactly when \( y_{it} \leq (1 - D)\mu_t \) which from Equation (2.4) is when \( I(f^{(-1)}(y_{i0}), \epsilon_{it}) \leq (1 - D)\mu_t \). Therefore in the \textit{ex post} case a voter’s redistributive preferences depend on the outcome of idiosyncratic shocks.

However, if voting takes place \textit{ex ante} before period \( t \), then voter \( i \) prefers \( r_1 \) to \( r_0 \) exactly when \( \mathrm{E}[y_{it}] = f^{(t)}(y_{i0}) \leq (1 - D)\mu_t \). Therefore all voters \( i \) of the same initial income level \( y_{i0} \) vote identically. We assume that voting takes place \textit{ex ante} over redistribution, with the unrealistic consequence that if \( f \) is known perfectly, voters with the same initial income \( y_{i0} \) vote identically.

The demand for redistribution in each period will clearly depend on the distribution of income each period. For our purposes the most important aspect of the income distribution is the \textit{distribution of expected time-} \( t \) \textit{income}, \( F_t(y) = \mathrm{Pr}(E[y_{it}] \leq y) \). Given a continuum of individuals \( i \) with incomes \( y_{i0} \) that follow a distribution \( F_0(y) = \mathrm{Pr}(y_{i0} \leq y) \) at time 0, we have

\[
F_t(y) = \mathrm{Pr}(E[y_{it}] \leq y) = F_0(f^{(-t)}(y))
\]

“One-shot” redistributive preferences are the proportion of the population who, sitting at time \( t - 1 \) and taking an \textit{ex ante} vote, would prefer \( r_1 \) to \( r_0 \) at period \( t \). Thus one-shot preferences would reflect repeated annual polling for policies which last one year. In this case, voter \( i \) will prefer \( r_1 \) to \( r_0 \) at time \( t \) if and only if their expected income at time \( t \), \( E[y_{it}] = f^{(t)}(y_{i0}) \) is less than the redistributed income at time \( t \), \((1 - D)\mu_t \). Therefore the income of an individual who is indifferent between \( r_1 \) and \( r_0 \) at time \( t \) is \((1 - D)\mu_t \) and the income of this indifferent individual at time 0 is \( f^{(-t)}((1 - D)\mu_t) \).

Finally, within our linear taxation and redistribution scheme, if a voter prefers \( r_1 \) to \( r_0 \) he or she prefers \( r_\tau \) to \( r_\tau' \) for any \( \tau > \tau' \). Thus the conditions which summarize redistributive preferences at

\[\text{Note that in principal the distribution of expected time-} t \text{ income could vary depending on the information available in different periods. Assumption 1 relieves us of these complications since the expected income of an individual at time } t \text{ becomes deterministic.}\]
period \( t - 1 \) are given by Equations (2.6-2.7).

\[ (2.6) \text{ Cutoffs for Redistributive Preference} = \begin{cases} \text{Redistribute,} & f(y_{it-1}) \leq (1 - D)\mu_t \\ \text{Laissez Faire,} & f(y_{it-1}) \geq (1 - D)\mu_t \end{cases} \]

\[ (2.7) \text{ Fraction of Population by Preference} = \begin{cases} \text{Redistribute,} & F_0(f^{(-t)}((1 - D)\mu_t)) \\ \text{Laissez Faire,} & 1 - F_0(f^{(-t)}((1 - D)\mu_t)) \end{cases} \]

Looking at Equations (2.6-2.7), of particular interest when considering “one-shot” redistribution is how \( f^{(-t)}(\mu_t) \) evolves over time as this dictates the proportion of the population who wants redistribution, \( F_0(f^{(-t)}(\mu_t)) \). Clearly \( F_0(f^{(-t)}(\mu_t)) \) depends on the initial distribution of income \( F_0 \) as well as the sequence \( \{\mu_t\} \) which is fixed by \( \mu_t = \int f(t)(y)dF_0(y) \). Since voters’ preferences evolve with \( f^{(-t)}(\mu_t) \), the outcome of an election depend not only on the current period of the vote, but the longevity of the policies presented to voters. For instance, if the demand for one period redistribution will follow an upward trend for the next decade, a ten year redistributive policy should garner more immediate support than a one or five year policy. This is because longer policies force voters to commit to what they want “most of the time” through aggregation of their income dynamics.

2.3. Forward Looking Demand for Redistribution. We now consider redistributive preferences for policies that last from period 0 through period \( T \). We assume voters have additively separable utility with discount rate \( \delta \). The discounted stream of income a voter \( i \) receives from period 0 to \( T \) depends on both initial income \( y_{i0} \) and a history of idiosyncratic shocks:

\[ (2.8) \text{ Discounted Income Stream} = \sum_{t=0}^{T} \delta^t y_{it} = \sum_{t=0}^{T} \delta^t I(f^{(t-1)}(y_{i0}), \epsilon_{it}) \]

Pocketbook voters sitting at period 0 consider redistributive schemes \( r_1 \) and \( r_0 \). As before it is a voter’s expected income which matters in this calculation. For convenience we introduce a function \( g^T(y) \) which is the expected discounted income of a voter with initial income \( y \). Following from Equation (2.8) and \( E_0[y_{it}] = f^{(t)}(y_{i0}) \), \( g^T \) is given by Equation (2.9)

\[ (2.9) g^T(y_{i0}) = E_0 \left[ \sum_{t=0}^{T} \delta^t y_{it} \mid y_{i0} \right] = \sum_{t=0}^{T} \delta^t f^{(t)}(y_{i0}) \]
For notational convenience we also denote \((g^T)^{-1}\) by \(g^{-T}\). Similarly, define \(\mu^T\) to be the discounted mean of population income over periods 0 to \(T\) given by

\[
\mu^T = \int E_0[\sum_{t=0}^{T} \delta^t y_i]dF_0(i) = \sum_{t=0}^{T} \delta^t \mu_t
\]

Redistributive preferences for a policy which takes effect in period 0 and continues through period \(T\) must take into account both discounting and the evolution of an individuals income over the period. An individual with income \(y_{i0}\) in period 0 receives a discounted income stream of \(g^T(y_{i0})\) and if complete redistribution is enacted would receive a discounted income stream of \((1-D)\mu^T\). Consequently, this individual prefers redistribution over periods 0 to \(T\) if and only if \(g^T(y_{i0}) \leq (1-D)\mu^T\). To summarize, redistributive preferences when considering periods 0 to \(T\) are given by Equations (2.10-2.11) which are analogous to Equations (2.6-2.7) for "one-shot" redistribution.

\[
\text{Cutoffs for Redistributive Preference } = \begin{cases} 
\text{Redistribute,} & g^T(y_{i0}) \leq (1-D)\mu^T \\
\text{Laissez Faire,} & g^T(y_{i0}) \geq (1-D)\mu^T
\end{cases}
\]

\[
\text{Fraction of Population by Preference } = \begin{cases} 
\text{Redistribute,} & F_0(g^{-T}((1-D)\mu^T)) \\
\text{Laissez Faire,} & 1 - F_0(g^{-T}((1-D)\mu^T))
\end{cases}
\]

Again of particular interest is how \(g^{-T}(\mu^T)\) and thus the fraction of the population who wants redistribution, \(F_0(g^{-T}(\mu^T))\) changes as longer periods of policy commitment are considered. As \(g^{-T}(\mu^T) = g^{-T}(\int g^T(y)dF_0)\), \(g^{-T}(\mu^T)\) is a generalized average of the terms \(f^{(-t)}(\mu_t)\) which determine the demand for redistribution in the "one-shot" case. Therefore the role of the initial income distribution \(F_0\) is much the same: it affects the demand for redistribution though \(\mu^T\) and small perturbations in \(F_0\) can give rise to large changes in the demand for redistribution in the presence of poverty traps, especially for large \(T\). In the next section we consider how the shape of \(f\) determines changes in the demand for redistribution.

3. Political Evolution with Forward-looking Voters under Full Information

The Solow model of neoclassical economic growth relies on an assumption of diminishing capital returns to hypothesize that poorer nations will tend to catch up over time, or converge, with the
incomes of richer nations. When transported to the individual or microeconomic level, the Solow assumptions imply a process of convergence among the population of a single country.

**Figure 3.1. POUM and NoPOUM Income Transitions**

![POUM Income Dynamics](image1)

![NoPOUM Income Dynamics](image2)

Figure 3.1(a) illustrates a typical income dynamic implied by accumulation under decreasing returns. Note that this concave transition process implies a unique long-term or steady state income level, $y^*$, at the point where $f_c(y)$ crosses the 45-degree line. Under this transition process, individuals who begin with incomes below the steady state level will converge towards it, while those who begin above the steady state level will drop back towards it. Note that this sort of concave income distribution process offers prospects of upward mobility (POUM) to voters whose initial income levels are less than the steady state income level.

This Prospect of Upward Mobility for the poor to achieve convergence with the population at large can serve to lessen preferences for redistribution. This mechanism allows Benabou and Ok (2001) to connect income dynamics and aversion to redistribution through POUM. Subsequent evidence for the income based approach to voting has been mixed, and its reach can be extended by modeling more general income dynamics.

---

For an early review of both the theoretical and empirical controversies, see (Romer, 1994). A more recent review with a theoretical emphasis is (Azariadis and Stachurski, 2005). Fong (2001) finds that variables reflecting a personal benefit from redistribution are insignificant in predicting redistributive preferences in the US. On the other hand, Checchi and Filippin (2004) conduct experiments, finding some support that the POUM reduces chosen taxation rates and that longer time horizons tend to decrease chosen rates under POUM. Beckman and Zheng (2003) also find tentative support for the POUM hypothesis using surveys administered to undergraduates (primarily business and economics majors). At the international level, Wong (2002) examines the GSS and World Values Survey for redistributive preferences and finds the expected signs across incomes, but find no evidence of the “tipping behavior” implied by median voter or POUM models and that current. Wong also finds that expected income indicators are small in magnitude in explaining redistributive preferences.
In contrast to Figure 3.1(a), individuals need not face uniformly decreasing returns in asset accumulation. The increasingly well developed theory of poverty traps suggests a number of mechanisms that can trap households at low living standards (see the review in Carter and Barrett [2006]. Central to all of these theories of poverty traps is exclusion from financial markets. Put differently, if households have access to loan markets and insurance instruments, then even when confronted by locally increasing returns to scale and risk, they can successfully engineer a strategy to obtain the assets needed to jump to a high level equilibrium. But absent access to those financial markets, households below a critical initial asset level will remain stuck in a low level, poverty trap equilibrium.

The result of such poverty trap models is Figure 3.1(b) which illustrates income transition dynamics with multiple steady states. The non-concave income transition function, $f_n(y)$, maps incomes in period $t$ into incomes in period $t+1$. As can be seen, this non-concave transition function has multiple crossings of the 45-degree line and thus admits multiple equilibria: $y^*_H$ is the high income steady state; $y^*_L$ is the low level steady state. Bifurcation occurs around the unstable point $y_b$. Households that being with incomes in excess of $y_b$ will tend toward the high level equilibrium while those that begin below this critical threshold will head towards the low level, poverty trap equilibrium, $y^*_L$. This implies Limited or No Prospect of Upward Mobility (NoPOUM) for voters below the threshold $y_b$. In contrast to an economy with a concave income transition function, economic polarization will occur and inequality can deepen if income transitions are governed by a non-concave function like $f_n(y)$.

This section defines a general family that incorporates both $f_c$ and $f_n$ types of income transitions and then derives a general set of results with political implications. We formalize a general class of income transitions, characterize its properties and show that the class is rich enough to approximate any increasing and continuous income dynamic with arbitrary precision. We then show the general class can generate rich patterns in the demand for redistribution, providing a theorem which

---

9 There is now a plethora of theory about why financial markets are often thin, missing and, or biased against low wealth agents. Banerjee and Newman (1994) provide an example of a debt based poverty trap which generates income dynamics. For a recent contribution, see Boucher et al. (2007).


11 Strictly speaking, this non-concave income transition function implies increasing polarization, not necessarily increasing inequality, as Esteban and Ray (1994) discuss.
shows how these new income transitions can create both increases and decreases in the demand for redistribution.

3.1. A General Class of Income Transitions. The income transitions we consider, typified by [3.1 b], are continuous functions which have finitely many stationary points \(SP(p_1, \ldots, p_I)\) as defined in Equation (3.1).

\[
SP(p_1, \ldots, p_I) \equiv \{\text{injective } f \text{ where } p_i \text{ are precisely the fixed points of } f\}
\]

Here “SP” stands for “Stationary Points,” in other words \(SP(p_1, \ldots, p_I)\) consists of all continuous, 1-1 income transition functions with \(f(p_i) = p_i\) for each \(i\) and no other stationary points. Of special relevance to income dynamics, poverty traps consist of stationary points which attract poor individuals over time unless they can escape the trap so that \(SP(p_1, \ldots, p_I)\) makes such points of attraction explicit. Since strictly concave functions have at most two stationary points, say \(p_1\) and \(p_2\), the \(SP(p_1, \ldots, p_I)\) definition allows for essentially all of the income transitions considered by Benabou and Ok. We define the \(SP\) class of income dynamics as all functions in \(SP(p_1, \ldots, p_I)\) for some points \(\{p_i\}\).

\(SP\) is a broad class including “undulating” income dynamics as in [3.1 b]). Such undulation requires a minimal amount of curvature to avoid infinitely many stationary points. Lemma 1 provides mild sufficient conditions under which an income transition is in \(SP\). Curvature can be provided by strictly concave or convex functions, perhaps pasted together as in Lemma 1(1). A more general condition is that the curvature of the income transition doesn’t vanish around the stationary points as given in Lemma 1(2). A third condition we provide is that the dynamic can be differentiated until the sign of the derivative doesn’t change as in Lemma 1(3).\(^{12}\)

**Lemma 1** (SP Sufficient Conditions). \(f\) is a finite stationary point income dynamic if \(f\) is continuous, bounded, injective and either:

1. Piecewise strictly concave or convex on pieces with length \(\geq \epsilon > 0\).
2. (Curvature) If \(f(p) = p\) and \(f'(p) = 1\) then \(f''(p) \neq 0\).
3. \(f\) is \(k \geq 2\) times continuously differentiable and the \(k\)th derivative is non-zero.

**Proof.** See Appendix. \(\square\)

\(^{12}\)It can also be shown for any \(f \in SP(p_1, \ldots, p_I)\) that if \(F\) has support on \([p_1, p_I]\) then both \(f^{(t)}\) and \(f^{(-t)}\) remain in \(SP(p_1, \ldots, p_I)\) for all times \(t\). This is desirable as many equations throughout the paper involve the terms \(f^{(t)}\) and \(f^{(-t)}\) so we would like these expressions to stay in \(SP(p_1, \ldots, p_I)\). A proof may be found in the Supplemental Appendix.
Of empirical interest is the fact that our class is extremely broad, in the following sense: any continuous, increasing income transition on a bounded set can be approximated arbitrarily well by a member of the $SP$ class, stated formally as Proposition 1. Econometrically, this means that any estimated income dynamic is observationally equivalent to a member of the $SP$ class.

**Proposition 1 (SP Approximation).** Let $f$ be any continuous, increasing income transition on a set $[0, I]$. For any $\epsilon > 0$ there is a $g \in SP(p_1, \ldots, p_I)$ with $|f(x) - g(x)| < \epsilon$ for all $x \in [0, I]$. 

**Proof.** See Appendix □

Proposition 1 shows that for practical purposes we may consider all income dynamics (including both POUM and NoPOUM dynamics) as members of our general class. We now turn to consideration of how POUM and NoPOUM income dynamics influence the demand for redistribution.

### 3.2. Demand for Redistribution

From Section 2, the fraction of the population who demands redistribution at time $t$ is given by $F_0(f^{(-t)}(\mu_t))$. It follows that the path of redistributive preferences, namely if they are increasing or decreasing is determined by $f^{(-t)}(\mu_t)$. In a POUM world, the concavity of $f$ implies the relationship $f^{(-t+1)}(\mu_{t+1}) \leq f^{(-t)}(\mu_t)$ in each period so the demand for redistribution is always decreasing. Similarly if voters are forward looking, the fraction of the population who wants redistribution $F_0(g^{-T}(\mu^T))$ monotonically decreases as the longevity of the policy considered increases. Therefore in a POUM world, the demand for redistribution decreases with time, both in the sense of “one-shot” evaluations each period and in terms of the time commitment to a particular redistributive policy. This is the type of behavior that Benabou and Ok set out to explain. We summarize these two salient aspects of redistributive dynamics in a POUM world in a Corollary.

**Corollary.** [POUM Dynamics] Let $F_0$ be absolutely continuous and suppose $f$ is concave. Then:

1. The demand for “one-shot” redistribution decreases over time.
2. The demand for redistribution over a $T$ period horizon decreases in $T$.

However, non-convexities in a general $SP$ world can break the POUM relationship leading to more complex redistributive dynamics. In contrast to a POUM world in which the path of $F_t(\mu_t)$ is decreasing, in a NoPOUM world this path can be increasing in the presence of non-convexities, and need not even be monotone as incomes evolve. Thus, the twist in a NoPOUM world versus a POUM world is more complex movement in the demand for redistribution. In the appendix we show that for a large class of income dynamics, the demand for redistribution for a given dynamic

---

13These statements follow respectively from Proposition 2 and Theorem 3 of Benabou and Ok (2001).
is increasing or decreasing depending on the initial distribution of income. We illustrate the point here with a concrete example.

Consider an income distribution \( F_0 \) composed of \( N \) groups with incomes \( Y_1 < Y_2 < \ldots < Y_N \) from constituencies which make up fractions \( p_1, p_2, \ldots, p_N \) of the population. Then \( \mu_t = \sum p_i \cdot f^{(t)}(Y_i) \) and so clearly depends on the initial distribution of income \( F \) interacting with the evolution of group incomes \( f^{(t)}(Y_i) \). Small changes in \( F \) can also cause large qualitative changes in the demand for redistribution over time. Now suppose \( 0 < Y_1 < Y_2 < Y_3 < \bar{Y} \) and \( p_1 = p_2 = p_3 = 1/3 \). Also suppose \( f \) has three fixed points \( Y_{\text{Trap}} < Y_{\text{Escape}} < \bar{Y} \) where for \( y < Y_{\text{Escape}} \), \( f^{(t)}(y) \rightarrow Y_{\text{Trap}} \) and for \( y > Y_{\text{Escape}} \), \( f^{(t)}(y) \rightarrow \bar{Y} \). In this case all groups are going to either \( Y_{\text{Trap}} \) or \( Y_{\text{Escape}} \). and to fix ideas, assume \( Y_1 < Y_{\text{Escape}} < Y_3 \) so the \( Y_1 \) group converges to \( Y_{\text{Trap}} \) while \( Y_3 \) converges to \( \bar{Y} \). Clearly \( Y_1 \) prefers the complete redistribution scheme \( r_1 \) while \( Y_3 \) prefers \( r_0 \). This leaves open the middle class of “swing voters” \( Y_2 \). If \( Y_2 > Y_{\text{Escape}} \) then the middle class eventually climbs the income ladder to \( \bar{Y} \) and joins the (now majoritarian) voting block of \( Y_3 \). Otherwise, if \( Y_2 < Y_{\text{Escape}} \) there is a thinning of the middle class and swing voters eventually join with \( Y_1 \), implying the median voter prefers redistribution. What is interesting about this example is the fragility of the eventual voting outcomes: a small income difference \( \delta \) in \( Y_2 \) can push \( Y_2 + \delta \) to be greater or less than \( Y_{\text{Escape}} \). This eventually results in a large fraction \( p_2 \) of swing voters to switch their vote as income evolves. Similar consequences can arise if voters lack perfect information about \( F_0 \) or \( f \) so that small changes in beliefs can give rise to large changes in redistributive preferences.

To better describe this complex process, we connect the changing demand for redistribution to the upper and lower envelopes of an income dynamic. Our characterization explains POUM as a special case. However the characterization shows that when income dynamics exhibit convexity around the mean of the income distribution then NoPOUM occurs and the demand for redistribution increases. The take away message is that unlike the POUM world, in a \( SP \) world the demand for redistribution is not a foregone conclusion and may manifest in volatile political patterns. This points to evaluating the relationship between income dynamics and political choices in light of the income dynamics voters face on a country specific basis.

Our characterization relies on the current location of mean income \( \mu_t \) relative to the income transition \( f \). We first define the upper and lower envelopes of \( f \), \( \mathcal{U} \) and \( \mathcal{L} \) as the envelopes created by placing lines which are above and below \( f \). These are illustrated in Figure 3.2(b). Formally, this
Figure 3.2. Upper and Lower Envelopes

\[ U \equiv \inf \{ h(x) : h \text{ is a line and } h \geq f \} \]

\[ \bar{U} \equiv \sup \{ h(x) : h \text{ is a line and } f \geq h \} \]

Clearly for each \( y \) we have \( \bar{U}(y) \geq f(y) \geq U(y) \) and necessarily \( \bar{U} \) is concave and \( U \) is convex. As special cases, when \( f \) is concave \( \bar{U} \) and \( f \) coincide whereas when \( f \) is convex \( U \) and \( f \) coincide. Therefore in a POUM world, \( \bar{U} = f \). We define the sets of incomes where \( f \) and \( \bar{U} \) exactly coincide as \( \bar{u} \) and similarly define \( u \) as incomes where \( f \) and \( U \) coincide. The relationship of \( \bar{u} \) and \( u \) to the path of redistributive preferences is given in Proposition 2:

**Proposition 2.** If \( \mu_t \in \bar{u} \) then the demand for redistribution decreases in period \( t \) relative to period \( t-1 \). Conversely, if \( \mu_t \in u \) then the demand for redistribution increases.

**Proof.** We will consider \( \mu_t \in \bar{u} \) as the other case is similar. We want to show that \( f^{(-t)}(\mu_t+1) \leq f^{-t}(\mu_t) \). This holds iff \( \mu_{t+1} \leq f(\mu_t) \) and by assumption \( \bar{U}(\mu_t) = f(\mu_t) \) so we want that \( \mu_{t+1} = \int f(t+1)dF_0 \leq \bar{U}(\mu_t) \). Since \( f \leq \bar{U} \) we know that \( \int f(t+1)dF_0 \leq \int \bar{U} \circ f(t)dF_0 \) and so it is enough that \( \int \bar{U} \circ f(t)dF_0 \leq \bar{U}(\mu_t) = \bar{U}(\int f(t)dF_0) \) which follows from Jensen’s inequality. \( \square \)

Proposition 2 says that if the mean income next period \( \mu_t \) lies in the region \( \bar{u} \) corresponding to the upper envelope, the demand for redistribution decreases. Conversely, if \( \mu_t \) lies in the lower envelope region, the demand for redistribution increases. In this sense, the upper and lower envelopes of \( f \) are natural definitions of Right and Left income transitions based on \( f \). This also highlights the differences between POUM and NoPOUM income dynamics. In a POUM world, \( f \) is concave and
therefore corresponds to $U$. This implies that any $\mu_t$ always lies in the upper envelope region so the demand for redistribution is always decreasing (See Figure 3.2a). In contrast a NoPOUM income dynamic has regions which coincide with both the upper and lower envelopes. Depending on where $\mu_t$ lies the next period, the demand for redistribution can either increase or decrease (See Figure 3.2b). In the next section we apply this insight about the curvature to actual income dynamics in Latin America.

4. Demand for Redistribution in Latin America

We have just seen from the prior section that political dynamics for forward looking voters will depend on both the income transition and the initial distribution of income $F_0$. This section asks if these two considerations can help us understand recent electoral dynamics in Latin America.$^{14}$

Based on income decile data from three Latin American national surveys we estimate simple income dynamics and evaluate the implications in light of the above results.$^{15}$ We emphasize here that the estimates are very rough, leaving details to the Appendix. The results for Chile and Peru are presented graphically in Figure 4.1. The dashed diagonal line is the 45 degree line representing the break-even points on the income transition function. A benchmark of five years was chosen for graphical purposes as this roughly corresponds to the presidential election cycle in the countries we consider. Vertical lines fixed by the mid-points between first year deciles divide the ten income bins used for the estimation. Finally, the solid line represents the estimated income dynamic ($\hat{f}$) over five years (in other words, $\hat{f}^{(5)}$) for each country. As can be seen, the estimated dynamics for Peru show convexity for much of the income distribution and therefore exhibit NoPOUM dynamics. In contrast, the estimated dynamics for Chile show at least some prospects for absolute if not relative, mobility for all deciles of the income distribution.

As shown above, the nature of political preferences by forward-looking voters will depend fundamentally on the shape of the income distribution dynamic. Figure 4.2 graphs the demand for redistribution in Chile and Peru for policies which last for various lengths of time. In order to calculate these outcomes, we begin with the estimated income dynamic $\hat{f}$ for each country and an estimate of the distribution of income $\hat{F}_0$ which is interpolated from the data. Computing the demand for

$^{14}$ See (Lora and Olivera, 2005) for a review of the empirical literature and estimates of the political impact of specific economic policies.

$^{15}$ The data comes from SEDLAC (CEDLAS and The World Bank) and includes both monetary and non-monetary income as well as private/public Transfers and Imputed rent.
redistribution across time and for varying policy lengths then follows the development above. Explicitly, we calculate the estimated analogue of Equation (2.11) by setting $\hat{g}^T(y_i0) \equiv \sum_{t=0}^{T} \delta^t \hat{f}(t)(y_i0)$ for $\delta = .95$ and solving Equation (4.1) for each period and a range of values for $T$:

\begin{equation}
\text{Fraction of Population by Preference} = \begin{cases} 
\text{Redistribute}, & \hat{F}_0(\hat{g}^{-T}((1 - D)\hat{\mu}^T)) \\
\text{Laissez Faire}, & 1 - \hat{F}_0(\hat{g}^{-T}((1 - D)\hat{\mu}^T))
\end{cases}
\end{equation}

Figure 4.2 therefore illustrates the relationship of estimated income dynamics to redistributive preferences. At present we assume that redistribution incurs no dead weight loss so that $D = 0$. We now connect the income dynamics of Figure 4.1 to the demand for redistribution in both countries at the national level, associating the demand for redistribution with the politically Left.

First, consider the “one-shot” demand for redistribution in each country, or rather the fraction of voters who would prefer redistribution for one year. Chile shows a fairly linear pattern, which implies fairly flat redistributive preferences over the period as in Figure 4.2(a). In contrast, Figure 4.1 shows fairly convex income dynamics in Peru which imply an increasing demand for redistribution as we see.
Second, consider the demand for redistribution as the longevity of the redistributive period increases. Again we see barely any change in Chile, as implied by a fairly linear income dynamic. The convex dynamics for Peru imply redistributive demand increases with policy longevity: future redistributive pressures are averaged in with immediate demand for redistribution.

**Figure 4.2. Demand for Redistribution in Time/Longevity Space**

![Graph showing demand for redistribution in time/longevity space for Chile and Peru](image)

Consistent with the theory, the leftward political trend in each country facing income dynamics of Figure 4.2 is supported by survey evidence of ideological polarization and shifting perceptions of household economic status. However, Figure 4.2 also shows that existing levels of inequality at the national level should easily have created majoritarian support for redistribution in all time periods and countries. Thus the leftward shift observed in countries was late in coming if voters were aware of the income dynamics they faced. To put this in perspective, we consider the level of dead weight taxation loss that would provide majoritarian support for laissez faire policies. The levels of dead weight loss that correspond to Figure 4.2 are 46-51% in Chile and 45-51% in Peru which are all exceedingly high in comparison to existing estimates of dead weight loss (See for instance Olken (2006)). This effectively rules out dead weight loss as the sole explanation of support for laissez faire.

16Time trends for these survey questions across the countries are depicted in the Supplemental Appendix.
We address this issue in the next sections by modeling voters as Bayesian learners who started with a POUM prior as advocated by the Washington Consensus in the early 1990’s but who have updated their beliefs to reflect their experience, creating a rapid shift in favor of redistribution.

5. Learning and Voting Under Imperfect Information

While the analysis here helps make sense of the recent course of politics in Latin American countries under general income dynamics, our model predicts extremely high levels of preference for redistribution, even by the mid-1990s. This finding thus throws into sharp relief the question as to why so many voters voted for largely laissez faire policies prior to the early part of this century. One explanation of course is that the political space was tightly constricted during that initial time period, or perhaps that voters were simply fooled and voted against their economic self-interest.17

Another, perhaps complementary, explanation is to recognize the intrinsic complexity of income distribution dynamics. Income dynamics and prospects of (no) upward mobility are complex and hard to understand, especially in economies that have altered their economic model. Our analysis so far has followed Benabou and Ok and assumed that voters know the true income transition function and use this knowledge to construct their forward looking income forecast and vote accordingly. We now generalize this approach and model the evolving political preferences of voters who optimally learn from their own experience and continually update their understanding or mental model of the income distribution process in their economy.

To keep this problem manageable, we assume voters face a known family of possible income transition functions \( I_\lambda(y, \epsilon) \) indexed by the parameter \( \lambda \).18 Analogous to our earlier analysis, for any value of \( \lambda \), we can define the expected income transition as

\[
f_\lambda(y) \equiv \int I_\lambda(y, \epsilon) dP(\epsilon)
\]

where as before the individual idiosyncratic income shock \( \epsilon_{it} \) has distribution \( P \). The family of expected income transitions is assumed to be bracketed by two extreme specifications, one representing a right perspective or vision of how the economy operates (\( \lambda = 1 \)) and the other a left perspective (\( \lambda = 0 \)). We refer to these specifications as “ideologies,” using this word to denote a model or understanding of how the world works. We assume that the any income transition function

17 For consideration of a host of additional possibilities, see Putterman (1996).
18 We assume that each member of this family conforms to the continuity and other assumptions listed in note 5 above.
can be expressed as a linear combination of the left \((f_L)\) and the right \((f_R)\) ideologies:

\[
\lambda(y) = (1 - \lambda)f_L(y) + \lambda f_R(y)
\]

(5.1)

At any point in time \(t\), the individual’s understanding of the economy can be represented by a probability distribution \(\pi_{it}(\lambda)\) over possible values of \(\lambda\) while the true value of \(\lambda\), labeled \(\lambda_0\), is unknown. The individual’s expected income in period \(t + 1\) thus becomes:

\[
E[y_{it+1}|y_{it}, \pi_{it}] = \int_0^1 E[y_{it+1}|y_{it}, \lambda]\pi_{it}(\lambda)d\lambda = \int_0^1 f_{\lambda}(y_{it})\pi_{it}(\lambda)d\lambda
\]

Note that this specification allows us to naturally describe someone with a left view of the world as putting a lot of probability density on low or left values of \(\lambda\), whereas a right view of the world would be described as having probability massed near right side of the spectrum or 1. While this specification of how voters predict their future income under incomplete information can be easily incorporated into our model of forward-looking voters, we turn first to consider how the critical new element, the voter’s probability distribution over \(\lambda\), is formed and evolves over time.

5.1. A Bayesian Model of Voter Learning. We assume each voter \(i\) begins with a prior distribution \(\pi_{i0}(\lambda)\) over possible values of \(\lambda\) while \(\lambda_0\), the true value of \(\lambda\), is the same for all voters but is unknown. We also assume that voters are “backward looking” and keep track of their idiosyncratic income histories \(H_{it} = \{y_{i0}, \ldots, y_{it}\}\). The history \(H_{it}\) is used to update beliefs each period to a posterior belief \(\pi_{it}(\lambda|H_{it})\) according to Bayes rule. In our context, we can think of \(\pi_{i0}(\lambda)\) as an initial ideology a voter has about the income transitions they face, while \(\pi_{it}(\lambda|H_{it})\) is the voter’s new ideology after \(t\) periods of learning the true income dynamic they face.

In order to make this learning process concrete, we now analyze it assuming an explicit structure of the transient income shocks and their relationship to income each period in Assumption 2.

Assumption 2 (Personal Dynamics). The income dynamic each voter faces is characterized by:

1. At the true value of \(\lambda\), incomes are \(y_{it+1} = f^{(t)}_{\lambda_0}(y_{i0})\epsilon_{it+1}\).
2. The shock \(\epsilon_{it+1}\) is distributed Uniform\((1 - \sigma, 1 + \sigma)\) for some \(\sigma \in (0, 1)\).
3. Voters know the value of \(\sigma\).

Under Assumption 2 if the true value of \(\lambda\), namely \(\lambda_0\), were known then \(E[y_{it+1}] = f^{(t)}_{\lambda_0}(y_{i0})\) since \(E[\epsilon_{it+1}] = 1\). This coincides exactly with the perfect information case. Therefore the income dynamic specified in Assumption 2 can be interpreted as receiving some random percentage of
the “true” income one should receive given \( \lambda \). Depending on the magnitude of \( \sigma \), the amount of fluctuation is large or small.

Now consider how voters update their beliefs under Assumption 2. If \( \lambda_0 \) is the true value of \( \lambda \) then since \( y_{it+1}/f_{\lambda_0}^{(t)}(y_{i0}) = \epsilon_{it+1} \) each voter knows that

\[
(5.2) \quad \left| y_{it+1} - f_{\lambda_0}^{(t)}(y_{i0}) \right| / f_{\lambda_0}^{(t)}(y_{i0}) \leq \sigma
\]

Equation (5.2) encapsulates the fact that a voter knows his observed income \( y_{it+1} \) must be within \( \sigma \% \) of expected income \( f_{\lambda_0}^{(t)}(y_{i0}) \) under \( \lambda_0 \). Therefore any \( \lambda \) for which \( \left| y_{it+1} - f_{\lambda}^{(t)}(y_{i0}) \right| / f_{\lambda}^{(t)}(y_{i0}) > \sigma \) cannot be the true value of \( \lambda \). Eliminating these impossible values of \( \lambda \) is exactly what Bayes rule dictates as the updating rule.

5.2. Demand for Redistribution under Imperfect Information. In general, neither \( f_L \) nor \( f_R \) need reflect reality but rather idealized versions of what Left and Right ideologues might represent in a manifesto. Clearly the relative strength of a voter’s belief in these world views influences their voting behavior. These beliefs are measured by \( \pi_{it}(\lambda | H_{it}) \) which connect ideology to voting. In order to emphasize the role of beliefs in deriving a voter’s expected income, we now illustrate the decisions of a pocketbook voter who is also a Bayesian learner.

If a voter knows the true value of \( \lambda \) in period 0, \( \lambda = \lambda_0 \) then expected income in period 1 is given by \( E[y_{i1}|y_{i0}, \lambda] = f_{\lambda_0}(y_{i0}) \). However, each voter does not know \( \lambda_0 \) with certainty but has a prior distribution \( \pi_{i0}(\lambda) \) over possible values of \( \lambda \). In this case, the expected income in period 1 of a voter with income \( y_{i0} \) is a weighted average of expected income given each possible value of \( \lambda \), namely \( E[y_{i1}|y_{i0}, \lambda] = f_{\lambda}(y_{i0}) \) weighted by \( \pi_{i0}(\lambda) \). Therefore expected income in period 1 is given by

\[
E[y_{i1}|y_{i0}, \pi_{i0}] = \int_0^1 E[y_{i1}|y_{i0}, \lambda] \pi_{i0}(\lambda) d\lambda = \int_0^1 f_{\lambda}(y_{i0}) \pi_{i0}(\lambda) d\lambda
\]

At the end of periods 1 to \( t \), a voter updates his prior \( \pi_{i0}(\lambda) \) to a posterior \( \pi_{it}(\lambda) \) using his new history \( H_{it} = \{y_{i0}, \ldots, y_{it}\} \). Therefore expected income in period \( t + 1 \) is given by

\[
(5.3) \quad E[y_{it+1}|H_{it}, \pi_{i0}] = \int_{\text{Expected Income} | \lambda} f_{\lambda}^{(t+1)}(y_{i0}) \pi_{it}(\lambda | H_{it}) d\lambda
\]
Equation (5.3) highlights the two dynamic factors which influence a voter’s beliefs about expected income. The first element is expected income given \( \lambda \) is the true state of the world and is deterministic as in the first part of this paper. The second element is a voter’s ideological beliefs which evolve as information is collected in the form of the idiosyncratic income history \( H_{it} \).

As in Section 3, each voter has a future expected income of \( E[y_{it+1}|H_{it}, \pi_{i0}] \) next period while he believes the mean income next period will be \( E[\mu_{t+1}|H_{it}, \pi_{i0}] = \int \int E[y_{jt+1}|y_{j0}, \lambda] \pi_{it}(\lambda) d\lambda dF(y_{j0}) \). Therefore after accounting for any dead weight loss \( D \), a voter will prefer \( r_1 \) to \( r_0 \) if and only if \( E[(1-D)\mu_{t+1}|H_{it}, \pi_{i0}] \geq E[y_{it+1}|H_{it}, \pi_{i0}] \) or rather if and only if

\[
(5.4) \quad \int E[(1-D)\mu_{t+1} - y_{it+1}|y_{i0}, \lambda] \pi_{it}(\lambda) d\lambda \geq 0
\]

Since \( E[(1-D)\mu_{t+1} - y_{it+1}|y_{i0}, \lambda] \) is the expected transfer to a voter under \( r_1 \) given \( \lambda \) is the true state of nature, we can interpret Equation (5.4) as saying \( r_1 \) is preferred to \( r_0 \) whenever the expected transfer is positive, given initial income and beliefs. Note that two voters with the same initial incomes \( y_0 \) need not have the same redistributive preferences: whether Equation (5.4) holds depends on each voter’s income history through their ideology \( \pi_{it}(\lambda) \). This implies the popularity of redistributive policies varies in a non-trivial way across initial incomes. Voter preferences conditional on their history \( H_{it} \) are summarized in Equation (5.5).

\[
(5.5) \quad \text{Redistributive Preferences } | H_{it} = \begin{cases} \text{Redistribute, } & \int E[(1-D)\mu_{t+1} - y_{it+1}|y_{i0}, \lambda] \pi_{it}(\lambda) d\lambda \geq 0 \\ \text{Laissez Faire, } & \int E[(1-D)\mu_{t+1} - y_{it+1}|y_{i0}, \lambda] \pi_{it}(\lambda) d\lambda \leq 0 \end{cases}
\]

Equation (5.5) allows us to make a clear connection from ideological beliefs to demand for redistribution through the following assumption:

**Assumption 3.** Increases in \( \lambda \) imply relative income position improves \( \left( \frac{d}{dx} f^{(t)}(y_0) \geq \frac{d}{dx} E[\mu_{t+1}|\lambda] \right) \) for all swing voters defined as \( y_0 \in \left[ f^{(-t)}(E[\mu_{t+1}|\lambda = 1]), f^{(-t)}(E[\mu_{t+1}|\lambda = 0]) \right] \).

This Assumption says that as \( \lambda \) increases, each voter believes his expected income \( f^{(t)}(y_0) \) increases relatively more than mean income \( E[\mu_{t+1}|\lambda] \). Furthermore, we only require this to hold for voters who might potentially change their vote: the votes of both destitute (\( f^{(t)}(y_0) < E[\mu_{t+1}|\lambda] \) for all \( \lambda \)) and well-to-do (\( f^{(t)}(y_0) > E[\mu_{t+1}|\lambda] \) for all \( \lambda \)) are unaffected by belief.

Assumption 3 implies the expected transfer \( E[\mu_{t+1} - y_{it+1}|y_{i0}, \lambda] \) is decreasing in \( \lambda \). It follows that for a voter \( j \) with beliefs \( \pi_j(\lambda) \) “to the Right” of a voter \( i \) with beliefs \( \pi_i(\lambda) \) that voter \( j \)
tends to prefer less redistribution than voter $i$. To make this precise, assume that $\pi_j$ stochastically dominates $\pi_i$ and $y_{j0} = y_{i0} = y_0$. Since $E[\mu_{t+1} - y_{jt+1}|y_{j0}, \lambda] = E[\mu_{t+1} - y_{it+1}|y_{i0}, \lambda]$ is decreasing in $\lambda$, the stochastic dominance of $\pi_j$ over $\pi_i$ implies

$$
\int E[\mu_{t+1} - y_{jt+1}|y_0, \lambda] \pi_{jt}(\lambda) d\lambda \leq \int E[\mu_{t+1} - y_{it+1}|y_0, \lambda] \pi_{it}(\lambda) d\lambda
$$

Therefore in the absence of dead weight loss $D$, voter $j$ prefers less redistribution than voter $i$. This result which connects ideological belief to redistributive demand continues to hold for any level of dead weight loss provided $f_R \geq f_L$ and is summarized as Proposition 3.

**Proposition 3.** Suppose $f_R \geq f_L$ and Assumption 3 holds. If voters $i$ and $j$ are identical except voter $j$'s belief $\pi_j$ stochastically dominate voter $i$'s belief $\pi_i$ then $j$ prefers less redistribution than $i$.

In this framework, one would expect that the speed of learning would be related to both the variability of income signals and the gap between left and right predictions for an individual’s future income position. These expectations imply a rich set of testable implications about the evolution of political preferences and voting that we hope to explore in future work.

6. **The Right Left Political Shift in in Latin America**

This section employs the model of forward-looking, Bayesian voters to analyze the striking right to left political shift observed across contemporary Latin America. To do this, we first provide an empirically-grounded approach for representing left and right political ideologies. Second, arguing that economic crisis of the 1980s put the left in disarray, we argue that at the time of the transitions voters adopted a POUM prior as the economic crisis of the 1980s left no credible alternative to the emergent pro-market model. Applying these assumptions to Peru, we show that voter learning over the course of a dozen years would be expected to generate up to a 30 percentage point shift in the fraction of the electorate preferring redistributive to free market policies.

6.1. **Empirical Approximation of Left and Right Ideologies.** In order to arrive at plausible left and right ideological models of income dynamics, we construct two functions ($f_R$ and $f_L$) that literally surround the true (estimated) income transition function that we denote as $\hat{f}(y)$. We begin by characterizing the right income transition model as one that offers greater prospects for upward and implies less demand for redistribution than does $\hat{f}$. For a given $f_R$, we then residually construct $f_L$ so that the true function can be expressed as a linear combination of the left and right ideologies as specified in equation (5.1) above.
While this approach implies an element of arbitrariness, we keep our modeling options fairly open by defining \( f_R \) using a continuum of transition functions \( g_\rho(y) \) indexed by the parameter \( \rho \). Successively higher values of \( \rho \) correspond to more exaggerated right ideologies that promise greater upward mobility and imply less demand for redistribution. Given our method for residually calculating the left ideology, higher values of \( \rho \) also imply greater ideological polarization in the sense that the left and right positions become more sharply differentiated.

The conditions which characterize any such \( g_\rho(y) \) are surprisingly sharp as provided in Proposition 4.

**Proposition 4.** For any class of income transitions \( g_\rho(y) \) indexed by \( \rho \), demand for redistribution decreases in \( \rho \) for all income distributions if and only if \( \frac{\partial}{\partial \rho} \frac{\partial}{\partial y} \ln \frac{\partial}{\partial y} g_\rho(y) \leq 0 \) provided:

1. \( \frac{\partial}{\partial y} g_\rho(y) \) is positive and bounded.
2. \( \frac{\partial^2}{\partial y^2} g_\rho(y) \) is non-zero and continuous.

**Proof.** See Appendix. \( \square \)

We now describe the construction of \( f_R \) and \( f_L \) from the empirical income transition \( \hat{f}(y) \). First derive \( \overline{f}(y) \), the upper envelope of \( \hat{f}(y) \) (which is necessarily concave) and subtract a line \( sy \) to arrive at the component \( \overline{f}(y) - sy \). The component \( \overline{f}(y) - sy \) is concave and therefore POUM, and is added to the upper envelope \( \overline{f} \) to arrive at \( f_R \). Weighting the component \( \overline{f}(y) - sy \) by a factor \( \rho \) and adding it to \( \overline{f} \) allows us to define \( f_R \) as in Proposition 4 as

\[
(6.1) \quad f_R(y) = \overline{f}(y) + \rho[\overline{f}(y) - sy]
\]

Applying Proposition 4 to Equation (6.1) we have

\[
\frac{\partial}{\partial \rho} \frac{\partial}{\partial y} \ln \frac{\partial}{\partial y} f_R(y) = \frac{\partial^2}{\partial y^2} (f_R(y) - s\overline{f}(y)) \frac{\partial}{\partial y} \ln \frac{\partial}{\partial y} (f_R(y) - s\overline{f}(y)) \leq 0
\]

Therefore our definition of \( f_R \) implies lower demand for redistribution as \( \rho \) increases. \( f_R \) is therefore a POUM dynamic, which grows increasingly POUM as \( \rho \) increases. As this suggests, \( f_R \) defined in Equation (6.1) implies that voters who find \( f_R \) credible demand less redistribution than under \( \hat{f} \).

In order to find \( f_L \), we fix the true value of \( \lambda \) to 1/2, implying \( \hat{f}(y) = \frac{1}{2} f_R(y) + \frac{1}{2} f_L(y) \). Therefore \( f_L \) is fixed by the Right transition \( f_R \) and the empirical transition \( \hat{f} \) as

\[
(6.2) \quad f_L(y) = 2\hat{f}(y) - \overline{f}(y) - \rho[\overline{f}(y) - sy]
\]

\( \rho \)The role of the line \( sy \) is to ensure the positivity of \( f_L \), and the second derivative of \( sy \) vanishes so its inclusion does not affect curvature.
\( f_L(y) \) is the empirical transition \( \hat{f}(y) \) plus a component \( [\hat{f}(y) - \bar{f}(y)] - \rho[\bar{f}(y) - sy] \) so that if \( \hat{f} \) is POUM then \( f_L(y) = \bar{f}(y) - \rho[\bar{f}(y) - sy] \), the mirror of \( f_R \). Voters finding \( f_L \) credible demand more redistribution than under \( \hat{f} \). Furthermore, this redistributive gap between \( f_R \) and \( f_L \) increases with \( \rho \), making \( \rho \) an index of the ideological polarization between Right and Left positions.

Figure 6.1 illustrates the application of this approach to the cases of Chile and Peru. Using the estimated income transition functions described in the appendix, we derive the right ideology assuming that \( \rho_{\text{Chile}} = 80 \) and \( \rho_{\text{Peru}} = 5 \) which are the largest values that imply \( f_L \) is positive.

6.2. Learning and Political Dynamics under a ‘TINA’ Prior. The final element needed to permit analysis of Latin American political dynamics is a specification of voters’ initial beliefs about the prospects, or lack thereof, for upward mobility at the beginning of the 1990s. To illustrate the implications of our model, we take seriously the then common observation that there was an exhaustion of credible political alternatives to a liberal economic regime. As Margaret Thatcher famously intoned: “TINA–there is no alternative” to free markets.\(^{21}\) While perhaps an exaggeration, Thatcher’s statements motivates what we call the TINA prior, meaning an initial set of beliefs,

\[^{20}\text{The } s \text{ parameters have been established by setting } s = \bar{f}(\bar{y})/\bar{y} \text{ where } \bar{y} \equiv 2 \times 10^{th} \text{ Income Decile of Base Year. We have also imposed an income cap that is 250\% of the top decile in order to avoid discussing income dynamics far out of the range of our data.}\]

\[^{21}\text{Margaret Thatcher Foundation (1984-85).}\]
\( \pi_{i0}(\lambda) \), that heavily weight the right perspective on the income process and its promise of upward mobility. In the numerical analysis that follows, we assume that all voters begin with the prior probability distribution illustrated in Figure 6.2.

**Figure 6.2. The TINA Prior**

![TINA Prior Graph](image)

With this TINA prior in hand, and the empirically grounded representations of left and right ideologies in Figure 6.1, we are now in a position to numerically simulate political dynamics in Chile and Peru. Under the assumption that income process is noisy with an idiosyncratic income shock parameter \( \sigma = 1/2 \), Figure 6.3 shows the simulated evolution of political preferences for Peru in the initial period (circa 1990), six years later and 12 years later.\(^{22}\) The vertical axis in each figure represents the fraction of the electorate that would prefer redistributive economic policies using the forward-looking perspective developed earlier. The solid line in each figure represents what political preferences would look like under the assumption that all voters know the true income transition function; this solid line thus represents the same information discussed in section 4 above.

The dashed line in the Figures show the political preferences for voters who begin with the TINA prior and then experientially update the distribution they hold over the critical income distribution parameter, \( \lambda \). Interestingly, under the assumptions made, the median, forward-looking voter would have initially voted against redistribution given a TINA and substantial prospects for

\(^{22}\)Note the precise implementation of simulations–how many agents, etc.
upward mobility. However, after six years of living and learning from the actual income distribution process, the median voter, and most voters in the lower 60% of the income would have favored redistributive policies. After a dozen years, the preferences of most voters approach those that would hold under full information.

**Figure 6.3. Demand for a 10 Year Redistributive Policy by Initial Income**

![Figure 6.3](image)

(A) Peru: Year 0  
(B) Peru: Year 6  
(C) Peru: Year 12

Figure 6.4 provides another look at the political dynamics implied by our model of forward-looking, Bayesian voters. The vertical axis now displays the fraction of the electorate at each point in time that is expected to vote for redistribution. As can be seen, over the 1998 to 2010 simulation period in Peru, the fraction voting for redistribution rises by some 16% points, again approaching the levels that would be expected under full information by 2010.

These sharp swings in policy preferences are of course driven by swing voters’ radical reevaluation of their prospects for upward mobility as they learn from the actual operation of the Peruvian economy. An interesting contrast to these results is provided by undertaking a similar exercise for the Chilean economy. The estimated Chilean income transition function is one that shows absolute upward income mobility for all classes, though not much relative improvement for the initially lower income deciles. While simulated preferences for redistribution in the Chilean case are strong, they remain quite stable over time, offering a vision of a much more stable politics in Chile than in a country with its polarizing income distribution process.
6.3. **Dead Weight Loss and Political Volatility.** Our final simulation exercise explores the impact of dead weight losses attributable to redistribution policies on political dynamics. For purposes of the numerical analysis, we assume a modest 10% dead weight loss is known to accompany redistributive programs. The dotted lines in Figures 6.3 and 6.4 illustrate the simulated political dynamics for this dead weight loss case.

Not surprisingly, the presence of a dead weight loss dampens support for redistribution, initially reducing political support by almost 20 percentage points. More surprising is the finding that dead weight losses actually increase political volatility in the case of Peru. A dozen years of learning by voters (again assumed to begin with the TINA prior) returns support to redistribution almost to the levels expected when dead weight losses are zero. In the particular case of the Peruvian simulation, this learning effect in the presence of dead weight losses creates an almost 30 percentage point swing in the fraction of the forward-looking electorate that prefers redistributive policies.

That dead weight loss \((D)\) increases support for laissez faire is clear, but the large increase in volatility is perhaps surprising. As analyzed in the appendix below, this volatility effect is explained by the asymmetric effect that \(D\) would have on a Right partisan with a strong belief in \(f_R\) in comparison to a Left partisan with a strong belief in \(f_L\). Increases in \(D\) attrit support for \(f_L\).
redistribution much faster for a Right partisan than for a Left partisan, creating a wider gulf to
cross as voters learn. As individuals learn and their beliefs move toward $f_L$, their sensitivity to dead
weight losses evaporates, further powering a large shift in support to redistributive policies.

7. Conclusion

Adopting the perspective that voters are forward-looking and pay attention to income dynamics,
not just their static place in the income distribution, this paper has explored the left-right-left shift
in the politics of Latin American countries over the last three or four decades. Two analytical
innovations are key to this exploration. The first is the generalization of earlier work on forward-
looking voters to model political preferences under general families of income distribution dynamics,
not just under concave dynamics that offer prospects of upward mobility. This generalization,
motivated by empirical evidence of polarizing, non-concave dynamics that offer no prospects of
upward mobility for segments of the population, shows that preferences for redistributive policies
may increase, not decrease over time when voters are forward looking. However, detailed analysis
of the case of Peru suggests that there would have been initially strong support for redistribution
had voters been fully informed about the nature of the income distribution dynamics, making it
extremely hard to account for the elections in Peru and elsewhere in Latin America in the 1990s
that brought more conservative parties and candidates to power.

This observation motivates this paper's second innovation, namely its modeling of voters as
Bayesian learners who update their understanding of income distribution dynamics based on their
own lived experience. Given that most voters in Peru (and elsewhere in Latin America where the
late 1980s and early 1990s saw a transition to a market economy) had little prior experience with
the new economic model, we assume that they initially adopted a prior probability distribution
that put substantial weight on a right wing ideological position that attached strong prospects for
upward mobility to the region's new economic model. Numerical simulation of political preferences
as voters received noisy draws from the true (estimated) income distribution process shows that a
numerically substantial shift from strong right political majority to a strong left political majority
over the course of about a dozen years. Somewhat surprisingly, simulated political volatility for
Peru is actually increased when the electorate believes that redistributive policies carry dead weight
losses. While there can certainly be no claim that these patterns are to be expected everywhere, the
modelling approach does offer new ways to think about political economy, especially in transition or other economies where voters’ prospects for upward mobility are largely initially unknown to them.

References


**Appendix A. Empirical Details**


A *Simple Class of Income Dynamics*. We assume income evolves as $y_{it+1} = f_\beta(y_{it})$ where $\beta$ are parameters of an increasing income transition $f_\beta$. For expositional purposes, we have chosen a particularly simple form for $f_\beta$ to capture the rates of income change within each income decile. We assume $f$ is continuous and piecewise linear on segments $[\gamma_{i-1}, \gamma_i)$ which correspond to the $i^{th}$ income decile, and for convenience we define $\gamma_0 \equiv 0$ and $\gamma_{10} \equiv \infty$. The slope of $f_\beta$ on each segment $[\gamma_{i-1}, \gamma_i)$ must be positive and so is defined as $e^{\beta_i}$ where $\beta_i$ is the $i^{th}$ coordinate of $\beta$. Therefore $\beta$ is simply a vector containing the rate of income growth for each decile. The equation for $f_\beta$ is given in Equation (A.1).

$$ f_\beta(x) \equiv \sum_{i=1}^{10} e^{\beta_i} (x - \gamma_{i-1}) + \sum_{j=1}^{i-1} e^{\beta_j} (\gamma_j - \gamma_{j-1}) 1_{[\gamma_{i-1}, \gamma_i)}(x) $$

**Econometric Structure.** We are provided income deciles $I_{dt_k}$ for periods $\{t_k\}$ and deciles $d$, where we assume for convenience that $I_{dt_k}$ are the median of each decile observed with log-normal$(0, \sigma)$ errors $\xi_{dt}$. Letting $F_0$ denote the cumulative distribution of income in period 0, our assumptions imply that the observed $I_{dt_k}$ are given by Equation (A.2).

$$ I_{dt_k} = f_\beta^{(t_k)}(x_d) \xi_{dt_k} $$
where \( x_d \equiv F_0^{-1}(0.1d - 0.05) \) is the median income of the \( d^{th} \) decile starting in period 0. Given Equation (A.1) we recover \( \beta \) through maximum likelihood.

Estimates. Before getting to the estimates, we wish to point out some caveats about our results. Since we are working with income deciles rather than micro-data, the estimates should not be taken too literally. In particular, due to the small number of observations and ten parameters, confidence bounds for the estimates would be essentially meaningless. On the positive side, our estimates do most likely capture far more relevant information about income dynamics than point estimates such as GINI or Polarization measures. We report our country wide estimates of \( \beta \) for each country in Table 1.

<table>
<thead>
<tr>
<th>Country</th>
<th>Slope of Income Transition in each Decile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>1.0369 1.0280 1.0302 1.0290 1.0334 1.0233 1.0303 1.0291 1.0250 1.0000</td>
</tr>
<tr>
<td>Peru</td>
<td>1.0399 0.9903 0.9970 0.9831 0.9898 0.9995 0.9855 0.9948 0.9913 1.0000</td>
</tr>
<tr>
<td>Venezuela</td>
<td>0.9031 1.0759 0.7250 1.2172 0.8410 1.0245 0.9665 0.9686 0.9762 1.0000</td>
</tr>
</tbody>
</table>

Appendix B. Proofs

Lemma (SP Sufficient Conditions). \( f \) is a finite stationary point income dynamic if \( f \) is continuous, bounded, injective and either:

1. Piecewise strictly concave or convex on pieces with length \( \geq \epsilon > 0 \).
2. (Curvature) If \( f(p) = p \) and \( f'(p) = 1 \) then \( f''(p) \neq 0 \).
3. \( f \) is \( k \geq 2 \) times continuously differentiable and the \( k^{th} \) derivative is non-zero.

Proof. Since \( f \) is assumed continuous and 1-1 the only thing we need to show in each claim is that \( f \) has finitely many fixed points. For a strictly concave or convex function this follows from the well known fact that a strictly concave or convex function can only cross any particular line, in particular \( y = x \), at most twice. Both claims we will use the following argument to rule out infinitely many fixed points: Suppose \( f \) has infinitely many fixed points \( \{x_i\} \). Then \( \{x_i\} \) must be contained in some compact set \( K \) since if \( \inf \{x_i\} \) or \( \sup \{x_i\} \) fails to exist then since each \( x_i \) is a fixed point, \( \inf \{f(x_i)\} = \inf \{x_i\} \) or \( \sup \{f(x_i)\} = \sup \{x_i\} \) fails to exist, contradicting the boundedness of \( f \).

Claim 1: Denote the pieces on which \( f \) is strictly concave or convex and which contain fixed points by \( [z_i, z_{i+1}] \) with \( i \in \mathbb{Z} \) where \( i < j \) implies \( z_i < z_j \). By the above argument, each \( [z_i, z_{i+1}] \) must intersect \( K \) which is bounded since it is compact, and therefore there is some \( M \in \mathbb{R} \) s.t.
Claim 2: We assume \( f \in C^2 \). Let \( X \equiv \{ p : f(p) = p \} \) be the collection of all fixed points of \( f \) and we wish to show that \( X \) is finite. We first claim that \( X \) is compact. Since \( X \subset K \) compact it is enough that \( X \) is closed. Suppose \( \{ x_n \} \subset X \) with \( x_n \rightarrow x \) and the continuity of \( f \) implies \( f(x_n) \rightarrow f(x) \). Since \( f(x_n) = x_n, x_n \rightarrow x \) and \( f(x_n) \rightarrow f(x) \) we have \( f(x) = x \) so \( x \in X \) and \( X \) is compact. Therefore if \( X \) has no limit points, \( X \) may contain only finitely many points so WLOG assume \( X \) has at least one limit point \( x \) with \( \{ x_n \} \subset X \) s.t. \( x_n \rightarrow x \).

By above \( f(x) = x \) and since \( f(x_n) = x_n \) we must have Equation (B.1).

\[
(B.1) \quad f'(x) = \lim_{n \to \infty} \frac{f(x_n) - f(x)}{x_n - x} = \lim_{n \to \infty} \frac{x_n - x}{x_n - x} = 1
\]

Since \( f(x) = x \) and equation (B.1) holds, we have \( f''(x) \neq 0 \) by assumption so \( f \in C^2 \) implies there exists a neighborhood \( V \) of \( x \) s.t. \( f'' \geq 0 \) on \( V \). Now \( \forall h \) s.t. \( x + h \in V \) we have for some \( \xi \in (x, x + h) \) the second order Taylor expansion (B.2).

\[
(B.2) \quad f(x + h) = f(x) + \frac{f'(x)}{1!}h + \frac{f''(\xi)}{2!}h^2 = x + h + \frac{f''(\xi)}{2}h^2
\]

It is clear from Equation (B.2) that \( x + h \) cannot be a fixed point of \( f \) for \( x + h \in V \) since by construction \( f''(\xi) \neq 0 \) for each \( h \). We conclude for each limit point \( x \) of \( X \) there is a neighborhood \( V \) containing \( x \) and no other fixed points, contradicting the assumption that \( x \) is a limit point of \( X \). Since \( X \) is compact, this implies \( X \) is finite as desired.

Claim 3: By a similar argument as in Claim 2, we can use a \( k^{th} \) order Taylor expansion to assign a neighborhood to each fixed point which contains no other fixed points so the compactness of the domain give finitely many fixed points. \( \square \)

Proposition. Let \( f \) be any continuous, increasing income transition on a set \([0, I]\). For any \( \epsilon > 0 \) there is a \( g \in SP(p_1, \ldots, p_I) \) with \( |f(x) - g(x)| < \epsilon \) for all \( x \in [0, I] \).

Proof. The result clearly holds for the special case \( f(x) = x \) by choosing \( g(x) \equiv x + \frac{1}{N} \) for \( \frac{1}{N} < \epsilon \). Now, for any continuous function \( g \) on \([0, I]\) let \( T(g) \equiv \sup_{y \in [0, x]} g(y) \). Clearly any such \( T(g) \) is
increasing and it is easy to check that $T(g)$ is continuous. Furthermore, for any $f$ which satisfies the hypothesis we claim $\sup_{y \in [0,1]} |f(x) - g(x)| \leq \epsilon$ implies $\sup_{y \in [0,1]} |f(x) - T \circ g(x)| \leq \epsilon$. In order to see this, suppose $|f(x) - T \circ g(x)| > \epsilon$ for some $x$ and clearly this requires $T \circ g(x) > g(x)$ which implies $T \circ g(x) > f(x)$ since otherwise $|f(x) - T \circ g(x)| = f(x) - T \circ g(x) < f(x) - g(x) \leq \epsilon$. Letting $\alpha \equiv \inf\{y : \ T \circ g(y) = T \circ g(x)\}$ we have $T \circ g(x) = T \circ g(\alpha)$ and $T \circ g(\alpha) = g(\alpha)$ by continuity of $g$ so $$g(\alpha) = T \circ g(x) > f(x) \geq f(\alpha)$$ where the last inequality follows from the fact that $f$ is increasing. Therefore $|g(\alpha) - f(\alpha)| \geq |f(x) - T \circ g(x)| > \epsilon$, contradicting $\sup_{y \in [0,1]} |f(x) - g(x)| \leq \epsilon$.

Now let any $\epsilon$ and $f$ which is not the identity be given. By the Weierstrass approximation theorem there exists a polynomial $p$ s.t. $\sup_{y \in [0,1]} |f(x) - p(x)| \leq \epsilon$ and by the above result this implies $\sup_{y \in [0,1]} |f(x) - T \circ p(x)| \leq \epsilon$. Now we will show that excepting the case that $p$ is the identity, $T \circ p \in SP(p_1, \ldots, p_I)$ which completes the proof (since $f$ is WLOG not the identity we may assume WLOG $p$ is not the identity for the approximation). As above, $T \circ p$ is continuous, increasing and is bounded on $[0,1]$ so we need only show $T \circ p$ has finitely many fixed points. Now let $F$ be the set of fixed points of $T \circ p$ and we will show $T \circ p$ cannot have more fixed points than $p$. For each $z \in F$ either $T \circ p(z) = p(z) = z$ or $T \circ p(z) = z > p(z)$. In the latter case, let $$\tilde{a} \equiv \inf\{x \in [0,1] : \ T \circ p(x) = z\} \quad \tilde{b} \equiv \sup\{x \in [0,1] : \ T \circ p(x) = z\}$$ so that $T \circ p$ is constant on $[\tilde{a}, \tilde{b}]$ and $z$ is the unique fixed point of $T \circ p$ on $[\tilde{a}, \tilde{b}]$. If $p$ has no fixed points on $[\tilde{a}, \tilde{b}]$ then $p(x) \geq x$ on $[\tilde{a}, \tilde{b}]$ so $T \circ p(\tilde{b}) > \tilde{b}$ or $T \circ p(\tilde{a}) < \tilde{a}$ and therefore $T \circ p$ has no fixed points on $[\tilde{a}, \tilde{b}]$. We conclude for each $z$ that there exists at least one fixed point of $p$. Finally, since $q(x) \equiv p(x) - x$ is again a polynomial which is not identically zero, $q$ has finitely many zeros so $p$ has finitely many fixed points. Therefore $T \circ p$ has finitely many fixed points and $T \circ p \in SP(p_1, \ldots, p_I)$ for some $\{p_i\}$. \hfill \Box

**Lemma.** Suppose $f$ and $g$ have bounded derivatives. Then $g$ is concave iff $g(\int f dF) \geq \int g \circ f dF$ for all distributions $F$.

**Proof.** If $g$ is concave the result holds by Jensen's inequality. For the converse, assume $g(\int f dF) \geq \int g \circ f dF$ for all distributions $F$ and fix $a, b \in \mathbb{R}$ and $\lambda \in [0,1]$. Define the distribution $H_\delta(x) \equiv \begin{cases} \frac{x - a}{\lambda b - a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$
evaluate (B.4)

Similarly, from Equations (B.4), for $R \parallel (B.5)$ where the last line follows from Equations (B.4). It follows that for $L \parallel (B.6)$.

Proof. Let $f_\rho(y) \equiv g(y, \rho)$ and we want to show for all $\Delta > 0$ that

$$f_\rho^{-1}\left(\int f_\rho dF\right) \geq f_\rho^{-1}\left(\int f_{\rho+\Delta} dF\right)$$

**Proposition.** For any class of income transitions $g(y, \rho)$ indexed by $\rho$, demand for redistribution decreases in $\rho$ for all income distributions if and only if $\frac{\partial}{\partial \rho} \ln \frac{\partial}{\partial y} g(y, \rho) \leq 0$ provided:

1. $g$ is strictly increasing in $y$ and twice continuously differentiable
2. $\frac{\partial}{\partial y} g(y, \rho)$ is bounded and $\frac{\partial^2}{\partial y^2} g(y, \rho)$ non-zero

Proof. Let $f_\rho(y) \equiv g(y, \rho)$ and we want to show for all $\Delta > 0$ that

$$f_\rho^{-1}\left(\int f_\rho dF\right) \geq f_\rho^{-1}\left(\int f_{\rho+\Delta} dF\right)$$
for all distributions $F$ iff $\frac{\partial}{\partial y} \ln \frac{\partial}{\partial y} f_\rho(x) \leq 0$. Letting $h$ be defined by $f_\rho + \Delta = h \circ f_\rho$ since $f_\rho$ is strictly increasing Equation (B.5) is equivalent to $\int f_\rho dF \geq h^{-1} (\int h \circ f_\rho dF)$ (that $h$ has an inverse follows from $g$ strictly increasing in $y$). It therefore follows from Lemma ?? that Equation (B.5) holds iff $h = f_\rho + \Delta \circ f_\rho^{-1}$ is concave. Noting that $h' \circ f_\rho = f_\rho' + \Delta / f_\rho'$ we have

$$h'' \circ f_\rho \cdot f_\rho' = \frac{[f_{\rho+\Delta}' f_\rho'' - f_{\rho+\Delta} f_\rho''']}{(f_\rho'')^2}$$

where $h''$ exists by inspection. We conclude $h$ is concave iff $h'' \leq 0$ iff $f_{\rho+\Delta}' / f_{\rho+\Delta} \leq f_\rho'' / f_\rho'$ iff

$\frac{\partial}{\partial \rho} \frac{\partial}{\partial y} \ln \frac{\partial}{\partial y} g(y, \rho) \leq 0$. 

\[ \square \]

B.1. Non-monotonicity of redistributive demand. In a NoPoUM world, the determination of whether the demand for redistribution is increasing or decreasing depends on all of $f, t$ and the initial distribution of income. Although the demand for redistribution may be directly computed, in general it is hard to derive a particular path analytically due to its dependence on the range of possible income distributions. In order to highlight this relationship, we provide an “Impossibility Result.” Our result shows for a fixed NoPoUM income dynamic that the demand for redistribution can be either increasing or decreasing depending on the income distribution. We state our result as

**Proposition 5.** Suppose $f \in SP(p_1, \ldots, p_I)$ and let $F$ denote a continuous distribution of income on $[p_1, p_I]$.

1. If $\mathcal{U} \cap [p_2, p_I]$ contains an open set there is an $F$ where the demand for redistribution always decreases.
2. If $\mathcal{U} \cap [p_1, p_{I-1}]$ contains an open set there is an $F$ where the demand for redistribution always increases.

**Proof.** See Supplemental Appendix. \[ \square \]

This proposition shows that a broad class of NoPoUM dynamics can exhibit either increasing or decreasing demand for redistribution. The deciding factor for redistributive dynamics even for a fixed NoPoUM dynamic $a$ is the initial distribution of income. This emphasizes the interrelationship between “Upward/Zero Mobility” in the dynamic role of income transitions and the “existing order” in the role of the income distribution: political implications cannot be drawn without considering both.
Appendix C. Dead weight Loss

In order to explain this effect, we depict idealized transitions $f_R$ and $f_L$ in Figure C.1. This figure supposes $f_R$ is concave as above while $f_L$ is convex (which is approximately true in applications). Fix any future mean income $\mu$ which is a fair margin above median income, and consider the level of support for redistribution next period as dead weight loss $D$ increases. Under $f_R$, the fraction of the population supporting redistribution falls from $F_0(f_R^{-1}(\mu))$ to $F_0(f_R^{-1}([1 - D]\mu))$ which in Figure C.1 is larger than the drop in support under $f_L$, namely $F_0(f_L^{-1}(\mu))$ to $F_0(f_L^{-1}([1 - D]\mu))$.

This asymmetric effect of dead weight loss holds because in the illustrated range, the concavity of $f_R$ implies $f_R$ is much flatter than $f_L$ which is convex. The precise conditions under which this argument apply are stated in Proposition C. Proposition C shows that modeling Right and Left ideologies reveals a second, new insight that dead weight loss can increase political volatility for forward looking voters.

**Figure C.1. Polarization from Dead weight Loss**

\begin{center}
\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure_c1.png}
\caption{Polarization from Dead weight Loss}
\end{figure}
\end{center}

**Proposition.** Assume $f'_R,f'_L > 0$, $f''_R < 0$, $f''_L > 0$ and $f_R(0) = f_L(0) = 0$ with $E[f_R(y)] \geq E[f_L(y)]$. If $f'_R(y_{equal}) = f'_L(y_{equal})$ then the difference in redistribution demanded between Left and Right increases for $D$ where $1 - D \geq f_R(z)/E[f_R(y_{equal})]$.

**Proof.** The difference in redistribution demanded between Left and Right is $f_L^{-1}([1 - D]E[f_L(y)]) - f_R^{-1}([1 - D]E[f_R(y)])$. We need to show that for all suitable values of $D$, $\frac{\partial}{\partial D} f_L^{-1}([1 - D]E[f_L(y)]) \geq 36$.
\frac{\partial}{\partial D} f_R^{-1}([1 - D]E[f_R(y)]). Evaluating both sides of the inequality yields

\frac{\partial}{\partial D} f_L^{-1}([1 - D]E[f_L(y)]) = -E[f_L(y)]/f'_L \circ f_L^{-1}([1 - D]E[f_L(y)])

\frac{\partial}{\partial D} f_R^{-1}([1 - D]E[f_R(y)]) = -E[f_R(y)]/f'_R \circ f_R^{-1}([1 - D]E[f_R(y)])

By assumption \(E[f_R(y)] \geq E[f_L(y)]\) and \(f'_R, f'_L > 0\) so it is sufficient to show

(C.1) \hspace{1cm} f'_R \circ f_R^{-1}([1 - D]E[f_R(y)]) \leq f'_L \circ f_L^{-1}([1 - D]E[f_L(y)])

Constructing a distribution \(G(y) \equiv [1 - D]F(y) + D1_{y \leq 0}\), Jensen’s inequality with \(f_R(0) = f_L(0) = 0\) implies

\[ f_R^{-1}([1 - D]E[f_R(y)]) \leq \int x dG = [1 - D]E[y] \leq f_L^{-1}([1 - D]E[f_L(y)]) \]

so with \(f''_L > 0\) to show (C.1) it is sufficient that

\[ f'_R \circ f_R^{-1}([1 - D]E[f_R(y)]) \leq f'_L \circ f_R^{-1}([1 - D]E[f_L(y)]) \]

which holds for all \(D\) with \(f_R^{-1}([1 - D]E[f_R(y)]) \geq y_{equal}\) giving the result. \(\square\)