THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 12/07

Optimal control of inequality under uncertainty

By

Martin Forster, Davide La Torre, Peter Lambert

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Optimal control of inequality under uncertainty

Martin Forster[*] Davide La Torre[†] Peter Lambert[‡]

March 6, 2012

Abstract

We consider the optimal control of inequality under uncertainty, with a particular focus on income inequality. For an economy experiencing economic growth and random shocks, we show how a simple loss and 'bequest' function may be combined to guide the expected level of inequality towards a pre-defined target within a finite planning horizon. Closed form solutions show that, the stronger the shocks to the income distribution, the more aggressive is policy. We discuss the results in the context of recent applied and policy literature on social inequality, globalisation and economic instability.

JEL codes: C61, D31, D63, I38

Keywords: Globalisation; Inequality; Stochastic dynamic programming

[*]Corresponding author. Department of Economics and Related Studies, University of York, York YO10 5DD, U.K.. e-mail: mf8@york.ac.uk, tel.: +44 (0)1904 323797.

[†]Department of Economics, Business and Statistics, University of Milan, 20122 Milan, Italy. e-mail: davide.latorre@unimi.it, tel.: +39 02 50321462.

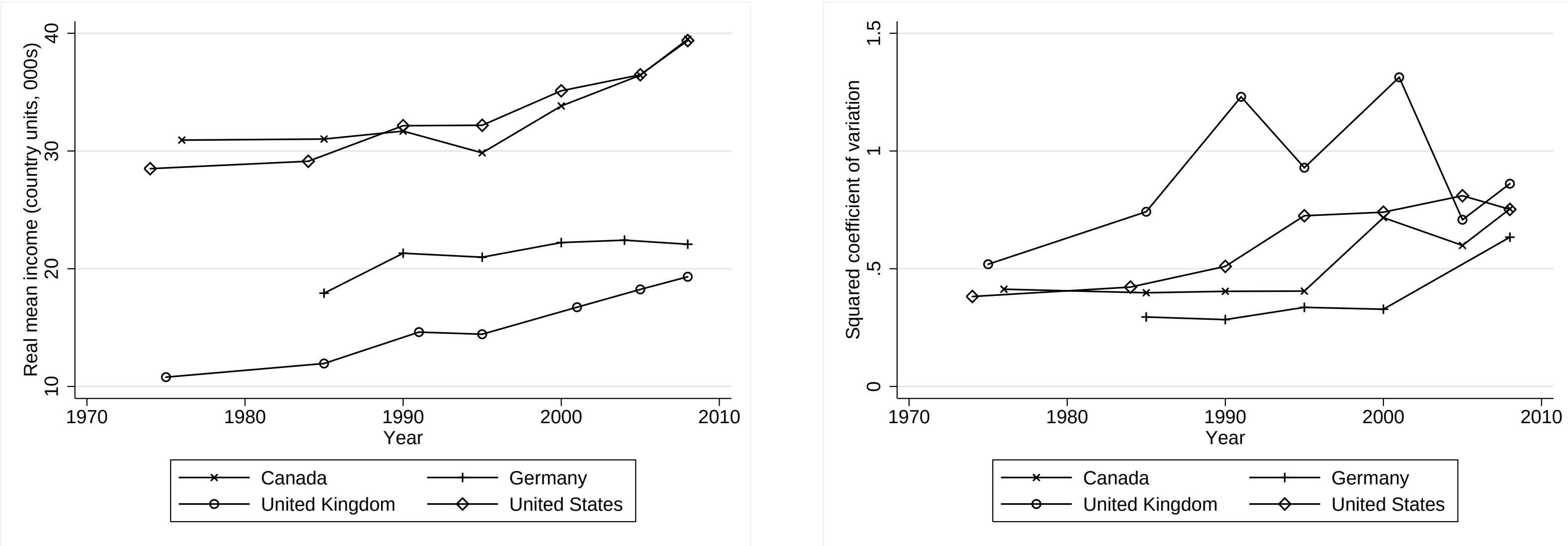
[‡]Department of Economics, University of Oregon, Eugene, OR 97403-1285, U.S.A., e-mail: plambert@uoregon.edu, tel.: +00 1 541 346-4670.

1 Introduction

This paper presents a dynamic, stochastic, model of the optimal control of inequality in the presence of economic growth and uncertainty, focusing, in particular, on the role played by shocks to the income distribution. We consider a policy maker who chooses the level of a mean-preserving control variable to influence the rate of change of an index of inequality and who incurs losses associated with both the instantaneous level of inequality and the strength of policy. The inequality index we choose to work with is the coefficient of variation. This has many desirable properties, key among which is that it assists derivation of a closed form solution for the optimal policy rule. Though the focus is on income, the framework is general and can be applied to any time-varying random variable whose distribution is subject to shocks.

The research is timely for a number of reasons. Firstly, there now exists a large body of evidence suggesting that many countries are experiencing economic growth accompanied by increasing income inequality. The OECD (2008) reports 'moderate but significant growth' in the gap between rich and poor in around three-quarters of its member countries over the last twenty to thirty years. Using its latest data, for the years 1975 to 2008, Figures 1(a) and (b) contrast the increasing real, average, disposable incomes of four of its member states (Canada, Germany, the United Kingdom and the United States) with changes in income inequality, as measured by the square of the coefficient of variation. Average incomes and inequality follow an upward trend for all four countries. Using a wider group of countries, the recent 'World of Work' reports of the International Institute for Labour Studies (2008, 2010) report a similar picture. Between 1990 and 2000, approximately two-thirds of the 85 countries reviewed (including those in Asia, the Pacific, Eastern Europe, the former USSR, the Middle East and North and Sub-Saharan Africa) experienced an increase in income inequality as measured by the Gini index. For the 44 countries for which data is available through to 2005, two-thirds experienced an increase in income inequality.

Secondly, survey evidence taken over the last decade suggests growing



(a) (b)

Figure 1: Trends in: (a) real mean disposable incomes (country currencies) and (b) squared coefficients of variation for Canada, Germany, the United Kingdom and the United States, 1975 to 2008 (Source `http://stats.oecd.org/Index.aspx?QueryId=26067`, accessed 5 March, 2012)

dissatisfaction with the way that nations are handling issues concerning inequality and poverty. In the recent Eurobarometer survey (TNS Opinion & Social, 2010), 62% of those asked felt that the way inequalities and poverty were addressed in their country was bad (31% felt that it was good), 38% felt that the situation had worsened in the last five years (11% felt that there had been an improvement) and 51% felt that there would be no change during the forthcoming year (12% felt that there would be an improvement). Using slightly older data for 23 countries, the World Values Survey (`www.worldvaluessurvey.org`) shows similar results, the index of tolerance of large income inequalities falling from 6.5 (survey of 1988-1993) to 5.4 (1999-2004), where a decline suggests increasing intolerance (International Institute for Labor Studies, 2008).

Thirdly, there is a growing body of literature linking high levels of inequality to political instability. Dutt and Mitra (2008) found empirical evidence to support the theoretical work of Acemoglu and Robinson (2000, 2001) suggesting that unequal societies are more likely to fluctuate into and out of democracy, as the elite in a democracy have an incentive to seize power so as to avoid redistributive policies, while the poor have an incentive to overthrow dictatorships in order to reassert redistributive policies. Dutt and Mitra found that inequality significantly exacerbates political instability, with the causal direction running from inequality to political instability, rather than the other way round. Similar results have been reported in the work of Muller and Seligson (1987) and Alesina and Perotti (1996).

Finally, despite large bodies of applied work documenting the changes in, and correlates of, income inequality over time (recent examples include Lundberg and Squire (2003), Jenkins and van Kerm (2006) and Heathcote et al. (2010)), we are not aware of any theoretical work which has addressed directly the problem set out in this paper.

The major contributions are as follows. The parsimonious set up of the model allows us to derive a closed form solution for the optimal policy rule to control inequality. The rule is independent of the starting level of inequality and, the more uncertain is the world in which the policy-maker operates, the more aggressive is optimal policy. The policy-maker can choose to target a particular reduction in the expected level of inequality over a finite time horizon by means of a penalty, or bequest, function. The optimal policy rule is nonlinear in the time remaining in the planning horizon and, under particular conditions, it can be optimal for the policy maker to allow inequality to increase over time. The results suggest that policy rules which ignore uncertainty arising from shocks to the income distribution can result in targets for inequality reduction being missed.

The paper is organised as follows. Section 2 presents the model, with its main results laid out in two sections. The first presents a simple micro-level description of inequality in an economy in which individual incomes follow a random walk with drift. This gives some theoretical underpinning to the story of Figures 1(a) and (b) - economic growth accompanied by increasing inequality - and introduces the coefficient of variation as our measure of inequality. The second section formulates and solves the policy-maker's optimisation problem using the material from the first. Section 3 discusses the results, the limitations of the model, and concludes.

2 Analysis

2.1 The model

A policy maker (henceforth PM) wishes to choose an optimal rule to control income inequality in an economy between year $t = 0$ and $t = T$, where t is an integer and T is finite. The PM seeks to hit a target level of inequality at T. Define Y_t as a continuous random variable denoting the incomes in the population. Assume that the inequality index of interest is the coefficient of variation, defined as $x_t = \sigma_{Y_t}/\mu_{Y_t}$, where $\sigma_{Y_t} > 0$ is the standard deviation of Y_t and $\mu_{Y_t} > 0$ is its mean.

Given Y_0, we assume that Y evolves according to the following first order stochastic difference equation:

$$Y_{t+1} = (1 + \alpha_t)Y_t + \epsilon_t, \quad t = 0, \ldots, T, \tag{1}$$

where α_t and ϵ_t are random variables with the (time-invariant) expected values $\mu_\alpha > 0$ and $\mu_\epsilon = 0$ and variances σ^2_α and σ^2_ϵ, respectively, together with covariance $\sigma_{\alpha\epsilon} = \rho_{\alpha\epsilon}\sigma_\alpha\sigma_\epsilon$, where $\rho_{\alpha\epsilon}$ is the correlation coefficient. Given the assumption about μ_α, it is to be expected that average income in the population is increasing over time. The following proposition describes the evolution of inequality over time.

Proposition 1. *Under the individual income growth process of Eq. (1), inequality is unambiguously increasing over time, that is, $x_{t+1} > x_t$, for all non-degenerate income distributions.*

All proofs are presented in the Appendix.

Proposition 1 is in the spirit of the results of earlier work by Eden (1980) and Deaton and Paxson (1994), which were based on a simple random walk. We introduce positive drift for consistency with the story told by Figures 1(a) and (b). The result in Proposition 1 is used to define the level of growth of μ_Y relative to σ_Y in the PM's optimisation problem, to which we now turn.

2.2 The policy maker's optimisation problem

Although t was an integer for Proposition 1, here we assume that it is continuous, to allow us to use the tools of stochastic calculus to solve the problem. The PM assumes that the following system describes the evolution of $\mu_Y(t)$ and $\sigma_Y(t)$:

$$\dot{\mu}_Y(t) = r\mu_Y(t), \quad \mu_Y(0) = \mu_{Y_0}, \tag{2a}$$

$$\dot{\sigma}_Y(t) = [1 - \gamma(t)]a\sigma_Y(t), \quad \sigma_Y(0) = \sigma_{Y_0}, \quad \forall t \in [0, T], \tag{2b}$$

where $r > 0$ and $a > 0$ are exogenous growth rates. $\gamma(t)$ is a mean preserving variable under the control of the PM which alters the rate of change of the standard deviation of the income distribution, while leaving the rate of change of mean income untouched (for example, γ could describe the extent of a mean-preserving reallocation of income from rich to poor). We impose the restriction $\gamma(t) \geq 0$ for all t, to rule out the scenario in which the PM actively seeks to increase the standard deviation of the income distribution. Hence the optimisation takes place on the set $\mathcal{S} = \{(x, \gamma) : x \geq 0, \gamma \geq 0\}$. Following Proposition 1, we assume that, in the absence of policy intervention ($\gamma(t) = 0$ for all t), inequality is unambiguously increasing with time (that is, $a > r$).

The coefficient of variation, $x(t) = \sigma_Y(t)/\mu_Y(t)$, is assumed to be subject to random shocks, reflecting uncertainty in either Eq. (2a), or (2b), or both. By differentiating $x(t)$ with respect to time, substituting in Eqs. (2a) and (2b) and adding exogenous, independently distributed Gaussian shocks scaled by $x(t)$, we obtain the controlled stochastic differential equation:

$$\frac{\mathrm{d}x(t)}{x(t)} = [(1-\gamma(t))a - r]\mathrm{d}t + \sigma_X \mathrm{d}W(t), \qquad x(0) = x_0 \equiv \sigma_{Y_0}/\mu_{Y_0}, \quad \forall t \in [0, T], \tag{3}$$

where $\mathrm{d}W(t) = Z(t)\sqrt{\mathrm{d}t}$, $Z(t) \sim N(0,1)$, is the increment of a Wiener process and σ_X is the variance parameter, such that when $\sigma_X = 0$ we have the case of no uncertainty. In the event of the PM choosing a constant level of γ for all t, which we shall call $\bar{\gamma}$, Eq. (3) has the analytical solution:

$$\begin{aligned} x(t) &= x_0 e^{\left((1-\bar{\gamma})a - r - \frac{\sigma_X^2}{2}\right)t + \sigma_X W(t)}, \qquad (4) \\ \text{so that :} & \\ E[x(t)] &= x_0 e^{((1-\bar{\gamma})a-r)t} \text{ and} \\ \mathrm{var}(x(t)) &= x_0^2 e^{2((1-\bar{\gamma})a-r)t}(e^{\sigma_X^2 t} - 1). \end{aligned}$$

$x(t)$ has a log-normal distribution such that, in the absence of policy intervention ($\bar{\gamma} = 0$) and making the assumption, from Proposition 1, that $a > r$, both the expected level and variance of inequality increase with time. When the PM intervenes, the greater is $\gamma(t)$, the stronger is the policy taken to reduce the expected rate of growth of inequality. However, a positive value of γ does not necessarily imply a reduction in the expected level of inequality;

this will only be the case when $\gamma > 1 - r/a$ (see Eq. (3)).

The PM wishes to choose an optimal policy rule for γ so as to minimise a performance criterion, defined as the expectation of the sum of the discounted integral of a loss function over the planning horizon, and a function which penalises the level of inequality remaining at T:

$$\mathbb{E}_0 \left[\int_0^T e^{-\rho t} \ell(x(t), \gamma(t)) \mathrm{d}t + \phi P[T, x(T)] \right], \tag{5}$$

The loss function ℓ is assumed to be of class $\mathcal{C}^{2,2}$ and is increasing and convex in each of its arguments, penalising deviations from perfect equality ($x(t) = 0$) and the strength of the control policy $\gamma(t)$. The penalty function P is assumed to be of class $\mathcal{C}^{1,2}$ and is increasing and convex in the level of inequality remaining at the end of the planning horizon, $x(T)$. $\phi > 0$ is a weight attached to the level of inequality remaining at T, such that $\phi = 0$ implies that no penalty is incurred; different values of ϕ allow the PM to target different levels of $x(T)$. $\mathbb{E}_0$ is the conditional expectation operator at $t = 0$ given an initial level of inequality, x_0. $\rho > 0$ is the discount rate. The minimisation takes place subject to Eq. (3) and its associated boundary conditions.

We restrict attention to the set $\mathcal{U}$ of admissable controls, that is, controls in $\mathcal{S}$ which lead to a finite expectation in Eq. (5). Define the value function as:

$$V(t, x) = \min_{\{\gamma(s) \in \mathcal{U}\}} \mathbb{E}_t \left[\int_t^T e^{-\rho(s-t)} \ell(x(s), \gamma(s)) ds + \phi P[T, x(T)] \right], \tag{6}$$

subject to Eq. (3), where $\mathbb{E}_t$ is the conditional expectation operator at t given that $x(t) = x$.

The following proposition shows that the simple value function $V(t, x) = e^{-\rho t}(0.5)A(t)[x(t)]^2$, where $A(t)$ may be determined, is associated with a version of a quadratic loss function with an interaction between $x(t)$ and $\gamma(t)$ and a simple penalty function which yield a closed-form solution for the optimal choice of $\gamma(t)$. We shall denote this as $\gamma_s^*(t)$, where the subscript s denotes the stochastic version of the model (we shall use the subscript d for the deterministic version).

Proposition 2. *For all non-degenerate income distributions and assuming the following value function belonging to the Generalised Entropy family of inequality indices:*[1]

$$V(t,x) = \frac{e^{-\rho t} A(t)[x(t)]^2}{2}, \tag{7}$$

there exist the following loss and penalty functions:

$$\ell(x(t),\gamma(t)) = \frac{[x(t)]^2(1+[\gamma(t)]^2)}{2} \quad \text{and} \tag{8a}$$

$$P[T,x(T)] = \frac{e^{-\rho T}[x(T)]^2}{2}, \tag{8b}$$

such that: (a) the optimal level of the control policy is as follows:

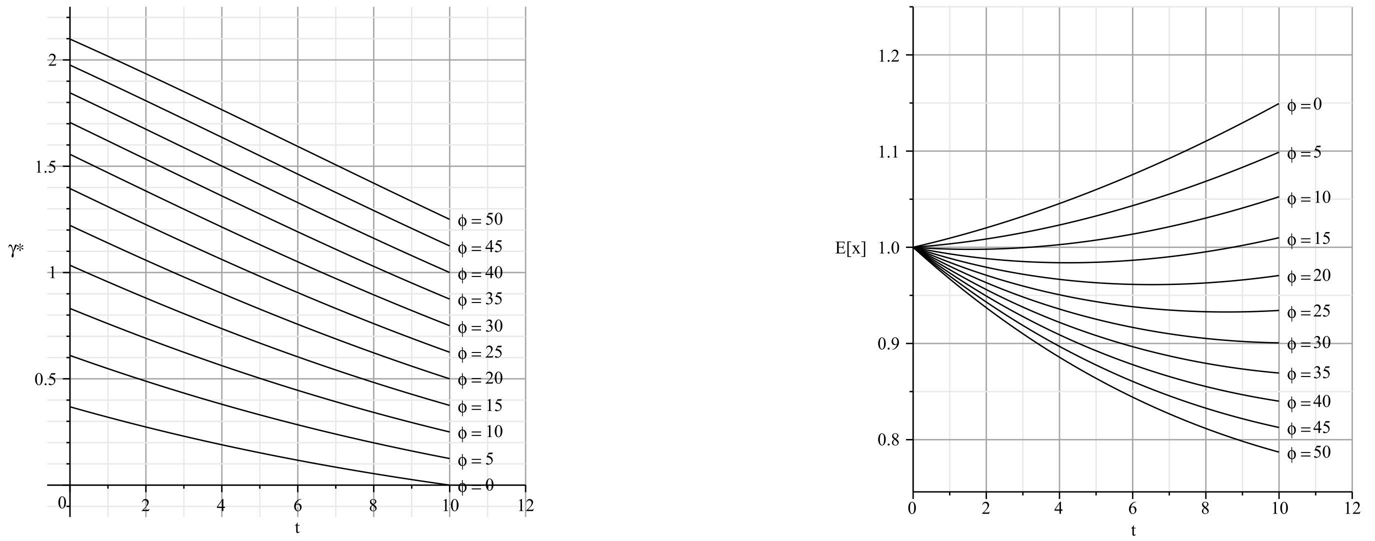
$$\gamma_s^*(t) = \frac{1}{2a}\Bigg(2(a-r) - \rho + {\sigma_X}^2 + \tanh\Bigg(\frac{\sqrt{M}(T-t)}{2} + \operatorname{arctanh}\left(\frac{2\,\phi\, a^2 + 2(r-a) + \rho + -{\sigma_X}^2}{\sqrt{M}}\right)\Bigg)\sqrt{M}\Bigg), \tag{9}$$

where $M = [2(a-r)-\rho+\sigma_X^2]^2+4a^2$, $\tanh(z) = (\exp(z)-\exp(-z))/(\exp(z)+\exp(-z))$ *is the hyperbolic tangent function and* $\operatorname{arctanh}(z) = (0.5)(\log(1+z) - \log(1-z))$ *its inverse; and (b) the optimal policy rule for control of income inequality is more aggressive in the stochastic version of the model than the deterministic version, that is,* $\gamma_s^*(t) \geq \gamma_d^*(t)$, *for all* $t \in [0,T]$.

[1] The Generalised Entropy family of inequality indices have the form:

$$\frac{1}{\alpha^2-\alpha}\int\left(\left(\frac{y}{\mu_Y}\right)^{\alpha} - 1\right) f_Y(y)\mathrm{d}y,$$

where α is real and not equal to zero or one. This expression equals the term in our value and loss function when $\alpha = 2$. Shorrocks (1980) developed these indices, which are the only measures for which relative inequality can be decomposed additively across population subgroups, a property which has found many uses; see Jenkins and van Kerm (2009) for a recent survey. Hence the value function in Proposition 2 is additively decomposable.



(a) (b)

Figure 2: Dependence of: (a) optimal policy rule γ_s^* and (b) optimal expected path of x on choice of ϕ

Eq.(9) shows that the optimal policy rule is independent of x_0 and is a nonlinear function of the time remaining in the DM's planning horizon, $T-t$. Conditional upon a particular choice of T, comparative static results for the other parameters in the model are difficult to establish in closed form, but differentiating Eq. (9) with respect to ρ, a and r and evaluating at a range of parameter values suggests that $\partial\gamma_s^*(t)/\partial\rho < 0$ (an increase in the rate of time preference makes optimal policy less aggressive), $\partial\gamma_s^*(t)/\partial a$ (the greater the rate of growth in the standard deviation of the income distribution, the more aggressive is optimal policy) and $\partial\gamma_s^*(t)/\partial r < 0$ (the greater the rate of growth of mean income, the less aggressive is optimal policy). These are intuitively agreeable results.

The remaining parameter in the expression for the optimal control is ϕ, the weight applied to the penalty function. Repeating the numerical comparative static analysis shows that $\partial\gamma_s^*(t)/\partial\phi > 0$, implying that, the greater the weight attached to the level of inequality remaining at T, the more aggressive is optimal policy. The effect of changing the level of ϕ on $\gamma_s^*(t)$ and the optimal expected path of $x(t)$ may be seen in Figures 2(a) and (b), where we run numerical simulations with the parameters $T = 10, a = 1/40, r = 7/1000, \rho = 1/50, \sigma_X = 1/4$, and $x_0 = 1$, varying ϕ. The higher is the value of ϕ, the more aggressive is the policy, and, consequently, the lower is the level of inequality remaining at T.

3 Discussion

Proposition 2 shows that a PM seeking to reduce inequality to a particular target level within a finite planning horizon should account for both the expected response of x to the policy variable and the strength of the shocks. Failure to account for the latter can lead to targets being missed, on average. This is an important result, given that many developed economies are currently experiencing sluggish growth and intensified economic instability, and are implementing wide-ranging austerity policies.

Proposition 2 also suggests a simple mechanism by which a PM may choose to target an expected level of inequality reduction by use of the weight

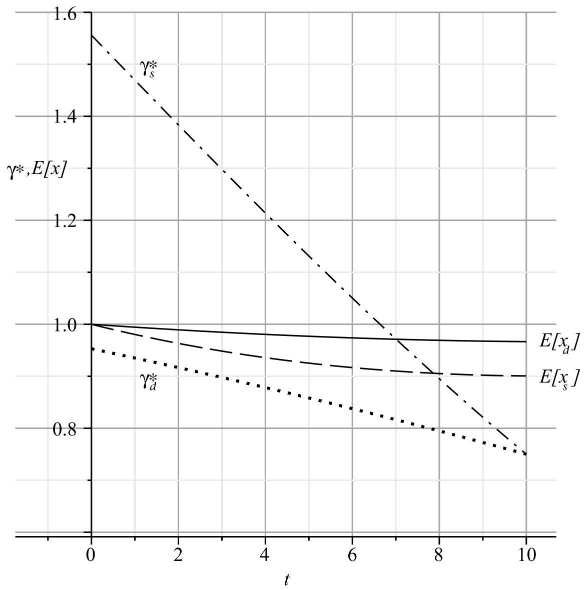


Figure 3: 'Missing the target', by failing to account for the shocks to the income distribution

ϕ attached to the penalty function. However, policy makers seeking to force the expected level of inequality to fall throughout the planning horizon must choose a value of ϕ which penalises sufficiently the level of inequality remaining at T (see Figures 2(a) and (b)). Figure 2(b) shows that the optimal expected path of x can be non-linear in t, meaning that, for a range of values of ϕ, the level of expected inequality can be increasing when following an optimal policy.

These ideas are illustrated in Figure 3, which contrasts the stochastic and deterministic policy rules for a PM seeking to reduce the coefficient of variation by 10%, from a starting value $x_0 = 1$, over a ten-year period and using the parameter values from the previous numerical example. We see from Figure 2(a) that this requires that $\phi \approx 30$. Figure 3 compares the resulting optimal paths of γ_s^* ($\sigma_X = 1/4$) and γ_d^* ($\sigma_X = 0$), and also shows the expected trajectories of x. It shows that failure to set a policy rule which accounts for the shocks to the income distribution, that is, setting the policy rule assuming that $\sigma_X = 0$, leads to the PM implementing a more benign policy, with the result that there is an expected decline in the coefficient of variation of 3.5% rather than the 10% achieved by γ_s^*.

As we pointed out in the introduction, the focus in this paper is on income inequality but the methodology can also be applied to other random variables. We have chosen to work with the coefficient of variation, a well-understood variability measure which is of particular interest in the measurement of income inequality (see on). Eden (1980) and Deaton and Paxson (1994) carry out most of their applied analysis using the variance of logarithms, which has some problems as an inequality measure (involving the principle of transfers, and enumerated in Foster and Ok, 1999), but they claim that their result, showing that inequality increases over time when income follows a random walk, holds for 'any measure of inequality that preserves the principle of transfers'. It is primarily through our choice of the coefficient of variation as an inequality index that we have been able to derive a closed form solution to the problem at hand.

There is a strong sense in which the coefficient of variation is salient for our model. Using the notation of section 2.1, define z as a variate which has been scaled to have the same mean, as well as the same Lorenz curve, as Y_{t+1} in expectation. We may think of z as the 'mean-and-inequality certainty-equivalent period 1 income distribution'. A necessary and sufficient condition for 'expected inequality' to have unambiguously increased according to any inequality index, is that the coefficient of variation of z exceeds that of income in period t.

There are many other possible candidates for inequality measure than the coefficient of variation in terms of which we have analyzed this problem. These include the ever-popular Gini coefficient. Further work could be done in terms of the Gini coefficient, although its non-differentiability in individual incomes would limit tractability. When the entire Lorenz curve for income is shifted up/down, every index of relative inequality shows an decrease/increase, and in such a case the choice of index is immaterial; but when there are Lorenz curve intersections, different indices respond differently. In many such cases the coefficient of variation is 'decisive', in that its directional change is reflected by other familiar inequality indices.[2] [3]

Our results show that ignoring the uncertainty associated with the growth equations for average income, the standard deviation of income, or both, could lead to policies missing their targets (in expectation). Of course, this has been demonstrated using a parsimonious set-up, with a mean-preserving control policy, which allowed us to derive a closed form solution. Extensions of the model in which, for example, the system of differential equations (2a) and (2b) is coupled, thereby allowing for feedback between the level of average income and inequality and vice versa (such as in the models of Lundberg and Squire (2003)) would be an interesting extension. In the present model, the 'costs' of policy are reflected in the loss function alone. A model which relaxes the assumption that the policy is mean preserving - for example, by incorporating a direct effect of government action on reducing the rate of growth of average income - might also reveal new insights.

Appendix

Proof of Proposition 1

We show the result for periods $t = 0$ and $t = 1$. The same argument may be used for subsequent periods. Using standard results for conditional means and variances (Wackerly et al., 2008), the coefficient of variation in period 1, x_1, is:

$$x_1 = \frac{\sqrt{\sigma_1^2}}{\mu_1} = \frac{\sqrt{(\sigma_{Y_0}^2 + \mu_{Y_0}^2)\sigma_\alpha^2 + 2\mu_{Y_0}\sigma_{\alpha\epsilon} + \sigma_\epsilon^2 + (1+\mu_\alpha)^2\sigma_{Y_0}^2}}{(1+\mu_\alpha)\mu_{Y_0}}. \tag{10}$$

[2] This result, which follows from Shorrocks and Foster's (1987) Corollary 1, encompasses all 'transfer-sensitive' inequality indices. Shorrocks and Foster argue that transfer sensitivity provides 'a means of prohibiting eccentric inequality judgements' (such as attaching greater importance to small transfers between millionaires than bigger transfers to the poor). Many of the generalized entropy indices are transfer sensitive.

[3] The Gini coefficient is not transfer sensitive, but it does satisfy a criterion called 'positional transfer sensitivity,' and it is similarly decisive for inequality comparisons using positionally transfer sensitive inequality indices when Lorenz curves cross once (Zoli, 1999). None of the positionally transfer sensitive inequality indices are differentiable functions of individual incomes.

Income inequality will rise, fall or stay the same between $t = 0$ and $t = 1$ according to:

$$x_1 \gtreqless x_0 \Leftrightarrow \sigma_{\alpha,e} \gtreqless -\frac{(\sigma_{Y_0}^2 + \mu_{Y_0}^2)\sigma_\alpha^2 + \sigma_\epsilon^2}{2\mu_{Y_0}}. \tag{11}$$

Since $\rho_{\alpha\epsilon} = \sigma_{\alpha,\epsilon}/\sigma_\alpha\sigma_\epsilon$, Eq. (11) may be written:

$$x_1 \gtreqless x_0 \Leftrightarrow -2\mu_{Y_0}\rho_{\alpha e} \lesseqgtr (\sigma_{Y_0}^2 + \mu_{Y_0}^2)\lambda + \frac{1}{\lambda} \equiv f(\lambda), \tag{12}$$

where $\lambda = \sigma_\alpha/\sigma_\epsilon$. It follows that $f'(\lambda) = (\sigma_{Y_0}^2 + \mu_{Y_0}^2) + 1/\lambda^2$, which has a unique minimum at $\lambda = \lambda^* = \frac{1}{\sqrt{\sigma_{Y_0}^2 + \mu_{Y_0}^2}}$. Hence:

$$f(\lambda) \geq f(\lambda^*) = 2\sqrt{\sigma_{Y_0}^2 + \mu_{Y_0}^2} = 2\mu_{Y_0}\sqrt{x_{Y_0}^2 + 1}, \quad \forall\lambda.$$

Returning to Eq. (12):

$$x_1 \gtreqless x_0 \Leftrightarrow -2\mu_{Y_0}\rho_{\alpha e} \lesseqgtr (\sigma_{Y_0}^2 + \mu_{Y_0}^2)\lambda + \frac{1}{\lambda} \geq 2\mu_{Y_0}\sqrt{x_{Y_0}^2 + 1}.$$

but since $-2\mu_{Y_0}\rho_{\alpha e} \in [-2\mu_{Y_0}, 2\mu_{Y_0}]$, it follows that $x_1 > x_0$ for all non-degenerate income distributions undergoing growth as defined by Eq.(1), that is, expected inequality in income, as measured by the coefficient of variation, is unambiguously higher in period 1 than in period 0. □

Proof of Proposition 2

The idea is to identify a $\mathcal{C}^{1,2}$ value function $V(t, x)$ which satisfies the Hamilton-Jacobi-Bellman (HJB) equation for the problem, together with the terminal condition given by the penalty function. This leads to the 'fundamental quadratic' for the problem, the solution to which can be used to establish the optimal policy rule. The HJB equation is:

$$-V_t = \min_{\gamma \in \mathcal{U}} \left\{ e^{-\rho t}\frac{x^2(1+\gamma^2)}{2} + V_x[(1-\gamma)a - r]x + \right. \tag{13}$$

$$\left.\frac{\sigma_X^2 x^2 V_{xx}}{2}\right\},$$

with terminal condition $V(T,x) = \phi P[T, x(T)]$. The optimal level of γ_s (a sufficiency condition is needed - see below) is obtained by solving Eq. (13):

$$\gamma_s^* = \frac{V_x a e^{\rho t}}{x}. \tag{14}$$

Substituting $\gamma = \gamma_s^*$ into Eq. (13) and simplifying gives:

$$-e^{\rho t} V_t = \frac{x^2}{2} - \frac{e^{2\rho t} V_x^2 a^2}{2} + e^{\rho t} V_x x(a-r) + \frac{e^{\rho t} \sigma_X^2 x^2 V_{xx}}{2}. \tag{15}$$

Make the guess that:

$$V(t,x) = \frac{e^{-\rho t} A(t) x^2}{2}, \tag{16}$$

where $A(t)$ is to be determined. This implies that:

$$V_t(t,x) = \frac{e^{-\rho t} x^2}{2}[\dot{A}(t) - \rho A(t)], \tag{17a}$$
$$V_x(t,x) = e^{-\rho t} A(t) x \quad \text{and} \tag{17b}$$
$$V_{xx}(t,x) = e^{-\rho t} A(t). \tag{17c}$$

By substituting Eq. (17b) into Eq. (14), we note that:

$$\gamma_s^* = a\,A(t). \tag{18}$$

Finally, substitute Eqs. (17a) to (17c) into Eq. (15) and cancel terms. We are left with the following ordinary differential equation in $A(t)$ (the RHS being the 'fundamental quadratic'), the solution to which may be used to yield the optimal γ (in Eq. (18)) and the optimal expected path of inequality:

$$\dot{A}(t,T) = -1 + (A\,(t))^2\, a^2 + 2\,A(t)(r-a) - \sigma^2 A\,(t) + A\,(t)\,\rho. \tag{19}$$

(a) The optimal level of γ_s^*

To obtain the optimal level of γ_s^*, solve Eq. (19) for $A(t)$ and substitute into Eq. (18) to give Eq. (9).[4] We use the stochastic maximum principle proposed by Framstad et al. (2004), to show that the policy rule Eq. (18) is optimal. Firstly, rewrite Eq. (13) as:

$$-V_t = \min_{\gamma \in \mathcal{U}} \{\mathcal{H}\},$$

where $\mathcal{H}$ is the stochastic Hamiltonian:

$$\mathcal{H} = e^{-\rho t}\frac{x^2(1+\gamma^2)}{2} + p(t)[(1-\gamma)a - r]x + \frac{\sigma_X^2 x^2 q(t)}{2}, \tag{20}$$

where $p(t)$ and $q(t)$ satisfy the adjoint equations for the problem. Theorem 2.1 of Framstad et al. states that, for an admissable set of state and controls, if $\mathcal{H}$ evaluated at the value of the control which minimises $\mathcal{H}$ is convex in x, for all t in $[0,T]$, then the pair (γ, x) comprise an optimal pair for the problem. $\mathcal{H}$ is strictly convex in γ since $\mathcal{H}_{\gamma\gamma} = x^2 e^{-\rho t} > 0$. The control which minimises H is given by Eq. (14) and so the minimised Hamiltonian is:

$$\hat{H}(x,\gamma,\psi(t),\pi(t)) = 1/2\, x^2\left(1 + \frac{p^2a^2}{\left(\mathrm{e}^{-\rho\, t}\right)^2 x^2}\right)\mathrm{e}^{-\rho\, t} + px \left(\left(1 - \frac{pa}{x\mathrm{e}^{-\rho\, t}}\right)a - r\right) + 1/2\, q\sigma^2 x^2$$

which is strictly convex in x, since $\hat{H}_{xx} = e^{-\rho t} + q\sigma^2 > 0$.

(b) Comparison of stochastic and deterministic policy rules

Setting $\sigma_X = 0$ in Eq. (9) does not allow a definitive comparison of the stochastic and deterministic policy rules, because σ_X appears in both the numerator and the denominator of the RHS. To compare the levels of γ_s^* and γ_d^*, we use a qualitative approach based on the analysis of the fundamental quadratic (following the ideas of Ewald and Wang (2011)).

Consider Eq. (19) when $\dot{A} = 0$ and let A_s^- and A_s^+ be the roots of the

[4] Maple 14 is used to solve Eq. (19).

fundamental quadratic (where $A_s^- < A_s^+$):

$$A_s^- = \frac{2(a - r) - \rho + {\sigma_X}^2 - \sqrt{M}}{2a^2}, \tag{21a}$$

$$A_s^+ = \frac{2(a - r) - \rho + {\sigma_X}^2 + \sqrt{M}}{2a^2}, \tag{21b}$$

where $M = [2(a-r) - \rho + \sigma_X^2]^2 + 4a^2 > 0$. It is straightforward to show that $A_s^- A_s^+ = -1/a^2 < 0$, implying that A_s^- is always negative and A_s^+ is always positive. Let A_d^- and A_d^+ be the roots when $\sigma_X = 0$. Inspection of Eqs. (21a) and (21b) shows that $A_d^- < A_s^-$ and $A_d^+ < A_s^+$.

The parabola of the fundamental quadratic has its vertex at $(A_s^{\min}, \dot{A}_s^{\min})$ where:

$$A_s^{\min} = \frac{2\,(a - r) - \rho + {\sigma_X}^2}{2\,a^2}, \quad \text{and} \tag{22a}$$

$$\dot{A}_s^{\min} = -\frac{M}{4a^2} < 0. \tag{22b}$$

Let $A_d^{\min}$ and $\dot{A}_d^{\min}$ be the respective values in the deterministic version of the model. By setting σ_X equal to zero in Eq. (22a), it follows that $A_d^{\min} < A_s^{\min}$ and $\dot{A}_d^{\min} > \dot{A}_s^{\min}$. By setting σ_X equal to zero in Eq. (22b), it follows that $\dot{A}_d^{\min} > \dot{A}_d^{\min}$.

Finally, for both the stochastic and deterministic versions, $\dot{A} = -1$ when $A = 0$.

Figure 4 contrasts the parabolas from stochastic and deterministic versions of the model under the restriction that $\rho < 2(a-r)$, which places the vertices of both parabolas to the right of the $\dot{A}(t)$ axis. The proof requires that the the parabola for the stochastic version of the model lies to the right of that for the deterministic version, which is guaranteed regardless of the value of the rate of time preference relative to a, r and σ_X^2 given that: 1. $A_s^- > A_d^-$ and 2. for both parabolas, $\dot{A} = -1$ when $A = 0$.

Use Eqs. (16) and (8b) to equate the value and penalty functions at T:

$$\frac{e^{-\rho T} A(T)[x(T)]^2}{2} = \frac{e^{-\rho T} \phi[x(T)]^2}{2}, \tag{23}$$

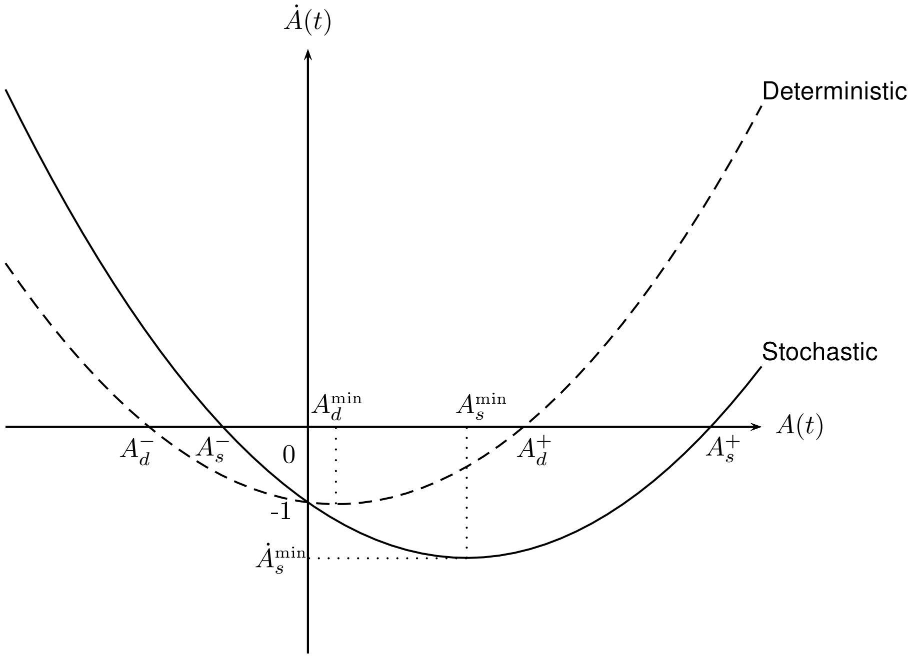


Figure 4: Phase diagram for stochastic and deterministic solutions to Eq. (19) under the restriction $\rho < 2(a - r)$

hence:

$$A(T) = \phi \geq 0, \tag{24}$$

which implies we may restrict attention to the orthants for which $A(t) > 0$.

The phase diagram shows that both A_d^+ and A_s^+ are unstable equilibria, because the quadratic is upward sloping at each root. $A(T) = \phi \geq 0$ may lie in three regions: $[0, A_d^+)$, $[A_d^+, A_s^+)$ and $[A_s^+, \infty)$. If $A_d^+ \leq \phi < A_s^+$, the path of $A(t)$ approaches $\phi = A(T)$ from below in the deterministic model and from above in the stochastic model. From Eq. (18), the optimal policy rule for the control is equal to $aA(t)$, so it follows that $\gamma_s^*(t) > \gamma_d^*(t)$ for all $t \in [0, T]$.

Now consider $0 < \phi < A_d^+$. The point $A(T) = \phi$ is approached from above by $A(t)$ in both the stochastic and deterministic models (because $\dot{A}(t) < 0$, refer to Figure 4). A simple contradiction may be used to show that it is still the case that, in this scenario, $\gamma_s^*(t) > \gamma_d^*(t)$ for all $t \in [0, T]$. Without loss of generality, fix a value of $A_s(T) = A_d(T) = \bar{\phi} > 0$ in this interval. We want to show that $A_s(t) \geq A_d(t)$ for all $t \in [0, T]$. Define the function:

$$g(t) = A_s(t) - A_d(t) \tag{25}$$

and let us suppose that there exists $t_1 \in [0, T]$ such that $A_s(t_1) - A_d(t_1) < 0$ (implying $\gamma_s^*(t_1) < \gamma_d^*(t_1)$). By applying a classical mean value theorem to Eq. (25), for some $t_2 \in [t_1, T]$:

$$\begin{aligned} \frac{[A_s(T) - A_d(T)] - [A_s(t_1) - A_d(t_1)]}{T - t_1} &= \dot{g}(t_2), \\ \Rightarrow \frac{-[A_s(t_1) - A_d(t_1)]}{T - t_1} &= \dot{g}(t_2) > 0, \end{aligned}$$

since $A_s(T) = A_d(T) = \phi$ and given our assumption that $A_s(t_1) - A_d(t_1) < 0$. This is a contradiction, since $\dot{g}(t) = \dot{A}_s(t) - \dot{A}_d(t) < 0$ for all $t \in [t_1, T]$ (refer to Figure 4). Hence $A_s(t) \geq A_d(t)$, and so $\gamma_s^*(t) \geq \gamma_d(t)$ by Eq. (18), for all $t \in [0, T]$.

Analogous reasoning can be used for $\phi \in [A_s^+, \infty)$. $\square$

References

Acemoglu, D. and Robinson, J. A. (2000). Why did the West extend the franchise? Democracy, inequality and growth in historical perspective. *Quarterly Journal of Economics*, 115(4):1167–99.

Acemoglu, D. and Robinson, J. A. (2001). A theory of political transitions. *American Economic Review*, 91(4):938–963.

Alesina, A. and Perotti, R. (1996). Income distribution, political instability and investment. *European Economic Review*, 40(6):1203–28.

Deaton, A. and Paxson, C. (1994). Intertemporal choice and inequality. *The Journal of Political Economy*, 102(3):437–467.

Dutt, P. and Mitra, D. (2008). Inequality and the instability of polity and policy. *The Economic Journal*, 118:1285–1314.

Eden, B. (1980). Stochastic dominance in human capital. *The Journal of Political Economy*, 88:135–145.

Ewald, C.-O. and Wang, W.-K. (2011). Analytic solutions for infinite horizon stochastic optimal control problems via finite horizon approximation: A practical guide. *Mathematical Social Sciences*, 61:146–151.

Foster, J. E. and Ok, E. A. (1999). Lorenz dominance and the variance of logarithms. *Econometrica*, 67:901–907.

Framstad, N. C., Oksendal, B., and Sulem, A. (2004). Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. *Journal of Optimization Theory and Applications*, 121(1):77–98.

Heathcote, J., Perri, F., and Violante, G. L. (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967 - 2006. *Review of Economic Dynamics*, 13(1):1094–2025.

International Institute for Labour Studies (2008). *World of Work Report 2008. Income Inequalities in the Age of Financial Globalization.* International Labour Organization (International Institute for Labour Studies), Geneva.

International Institute for Labour Studies (2010). *World of Work Report 2010. From one crisis to the next?* International Labour Organization (International Institute for Labour Studies), Geneva.

Jenkins, S. and van Kerm, P. (2009). *The Measurement of Economic Inequality.* Chapter 3 in W. Salverda, B. Nolan and T. Smeeding (eds.), The Oxford Handbook of Economic Inequality, Oxford University Press, first edition.

Jenkins, S. P. and van Kerm, P. (2006). Trends in income inequality, pro-poor income growth, and income mobility. *Oxford Economic Papers*, 58:531–548.

Lundberg, M. and Squire, L. (2003). The simultaneous evolution of growth and inequality. *The Economic Journal*, 113(487):326–344.

Muller, E. N. and Seligson, M. A. (1987). Inequality and insurgency. *American Political Science Review*, 81(2):425–452.

OECD (2008). *Growing unequal? Income Distribution and Povery in OECD Countries.* OECD Publishing.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48:613–625.

Shorrocks, A. F. and Foster, J. (1987). Transfer sensitive inequality measures. *Review of Economic Studies*, 54:485–497.

TNS Opinion & Social (2010). *Special Eurobarometer 315. Social Climate. Full Report.* European Commission.

Wackerly, D. D., Mendenhall, W., and Schaeffer, R. L. (2008). *Mathematical Statistics.* Brooks/Cole.

Zoli, C. (1999). Intersecting generalized Lorenz curves and the Gini index. *Social Choice and Welfare*, 16:183–196.