Three Essays on Environmental Economics

By

Jamie Donghoon Lee

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Economics

Guelph, Ontario, Canada

©Jamie Donghoon Lee, 2017
Abstract

Three Essays on Environmental Economics

Jamie Donghoon Lee

Advisors: Dr. Ross McKitrick and Dr. Thanasis Stengos

Three chapters are presented in this thesis, each in the field of environmental economics. The first chapter studies voting outcomes when people differ in their private expectations about marginal damages and the policy maker proposes an externality pricing instrument that is either based on a static political compromise or on a state-contingent updating rule (McKitrick, 2010). We examine cases in which voters are honest (they prefer the socially-optimal price based on their expectation of marginal damages) or dishonest (they prefer either a zero tax or a maximum tax on a priori grounds, irrespective of marginal damages) or combinations of these two. We show that when all voters are honest a standard pricing mechanism that aims at minimizing political losses may never obtain majority support, but implementation of the state-contingent pricing rule always obtains majority support. We then examine whether dishonest voters would prefer implementation based on the static rule or the state-contingent rule.

The second chapter examines the relationship between the social discount rate (a rate chosen by a social planner) and the pure rate of time preference (a rate chosen by a representative agent). We accomplish this by examining the slope of famous 'Ramey equation' with respect to the pure rate of time preference. In the literature, it is typically
assumed the value of the slope to be equal to one, which is equivalent to treating the consumption growth rate to be independent of the pure rate of time preference. Our general equilibrium framework shows that this relationship is rather complex, and depends on the relative signs and sizes of elasticities of consumption and the growth rate of consumption with respect to the pure rate of time preference. We then examine a closed form solution case where we observe their relationship depends on exogenous parameters of the model, such as initial capital stock, technology parameter as well as time index. Finally, we empirical estimate the relationship by using the Time-Varying Semiparametric Smooth Coefficient model using U.S data from 1930-2014. We find that the estimate of the slope of the social discount rate with respect to the pure rate of time preference to be 0.84 for the United States.

The social discount rate based on Ramsey equation is critically dependent upon future consumption growth rates and the pure rate of time preference. In practice, however, the pure rate of time preference is not observable, unlike with the consumption growth where past rates are known with certainty and future projections can be formed with some uncertainty. Ironically, numerous papers have incorporated and examined the effects of the consumption uncertainty but treat the pure rate of time preference to be constant and certain in social discount rates. The third chapter investigates the latter issue for various underlying cases in restricted and unrestricted frameworks. Overall, we find that the effects of uncertainty on the certainty-equivalent discount rate are critically dependent on the source of uncertainty and its assumed properties.
I am grateful to my supervisors, Dr. Ross McKitrick and Dr. Thanasis Stengos, for their great guidance, encouragement, suggestions, invaluable comments and infinite patience. Their contributions have made the completion of this thesis possible. I am here today, because of them. I would also like to thank a committee member, Dr. Chris McKenna for his guidance and encouragement.

Several other faculty members helped me though out my years as a student here at Guelph, whom contributed to my growth as a researcher. I would like to thank the following individuals: Dr. Maynard, Dr. Mei Li, Dr. Rene Kirkegaard and Dr. Johanna Goertz. I would also like to thank Dr. Richard Newell and Dr. William Pizer for sharing their data used in their paper 2003 paper. The data set contributed to the empirical component of chapter 2 of the thesis.

I thank my parents, my sister, Jenny and my brother-in-law Geoff, and my little niece, Amelia for their full support and bottomless love. To my lovely wife, Amy Shi, I love you and dedicate this thesis to you. I know without you, this thesis would have been impossible.

Last but not least, I would like to thank my friend, my brother, Tasos Papanastasiou, for his great friendship and love. Finally I would like to thank the entire faculty, staff members, and all graduate students in the Department of Economics and Finance. Great memories will never be forgotten.
# Table of Contents

List of Tables ........................................................................................................ vii

List of Figures ......................................................................................................... viii

1   **Forming a Majority Coalition for Carbon Taxes Under a State-Contingent Pricing Mechanism**

1.1 Introduction ........................................................................................................ 1

1.2 Voting on taxes for externalities and public goods ........................................... 5

1.3 The voting model ............................................................................................... 7

1.3.1 Voting on the simple static case ................................................................... 7

1.3.2 A simple static case ....................................................................................... 8

1.3.3 The policy maker’s problem ......................................................................... 9

1.3.4 The relationship between the variance of $\tau_i$ and voting outcomes......... 10

1.4 Voting outcomes when some of voters are honest and others are not .......... 15

1.4.1 The current problem with static CO2 emission tax policy ......................... 15

1.4.2 The loss function under the SCPM .............................................................. 15

1.4.3 Voting outcome when all voters are honest ................................................. 16

1.4.4 Voting outcomes when some or all voters are dishonest ........................... 18

1.5 Conclusions ...................................................................................................... 25

2   **Discounting when The Consumption Path Is Affected by the Rate of Time Preference**

2.1 Introduction ....................................................................................................... 27

2.2 Past theoretical and empirical works on the social discount rates ................ 30

2.3 The sign of the slope in the general equilibrium model .................................. 36

2.3.1 Derivation of Ramsey equation .................................................................. 36

2.3.2 Endogenous response to changes in $\rho$ and the general equilibrium value of $r$ ... 38

2.3.3 A closed form solution model ..................................................................... 40

2.4 Empirical estimates of $\rho$, $\eta$ and the slope for the United States ............... 46

2.4.1 Time-varying semi-parametric smooth coefficient model ....................... 46
2.4.2 Data.........................................................................................................................49
2.4.3 Results and Discussion ............................................................................................52
  2.4.3.1 Estimation results ..................................................................................................52
  2.4.3.2 Estimation of the slope .......................................................................................53
2.5 Conclusions ..................................................................................................................56

3 Discounting when The Consumption Growth Path and The Rate of Time Preference are Uncertain and Latter Affects the Former: A Simulation Approach

3.1 Introduction ..................................................................................................................58
3.2 Literature review .........................................................................................................60
3.3 Uncertainties in the consumption growth rates and the pure rate of time preference: The restricted approach ........................................................................................................63
3.4 Properties of \( \rho(H) \)..................................................................................................64
  3.4.1 Properties of permanent component \( \rho(H) \) hat in \( \rho(H) \)..........................65
  3.4.2 Uncertainty (random component) in \( \rho(H) \) ...................................................66
  3.4.3 General properties of certainty-equivalent \( r(t) \) with respect to uncertainties in the consumption growth rate and \( \rho \) ..........................................................................................................................67
3.5 Monte-Carlo simulations for restricted framework ....................................................69
  3.5.1 Step-by-step Monte-Carlo simulation procedure .................................................69
  3.5.2 Results ..................................................................................................................69
3.6 The pure rate of time preference dependent consumption growth path and the certainty-equivalent \( r(t) \): The unrestricted approach ........................................................................73
  3.6.1 A closed form model .............................................................................................73
  3.6.2 Monte-Carlo simulation steps for a closed form solution model .......................75
  3.6.3 Results of a closed form model ............................................................................75
3.7 Conclusions ..................................................................................................................78
References .........................................................................................................................80
List of Tables

2.1 Descriptive summary statistics of consumption growth and real interest rates for U.S. from 1930 to 2014 ..........................................................51
2.2 Descriptive summary statistics of the slope for U.S. from 1930 to 2014 .........................54
3.1 Descriptions of cases in Figure 3.2 (cases 1-3) and Figure 3.3 (cases 4-7) .....................70
3.2 Descriptions of cases in Figure 3.4 (cases 8-10) and Figure 3.5 (cases 11-13) .................75
List of Figures

1.1 Case where $V' > \text{tao bar}$ .................................................................21
2.1 Optimal capital stock over time for three values of $\rho$ .................................43
2.2 Optimal values of $r(t)$ over time for three values of $\rho$ ..................................43
2.3 Equation (34) (Marginal effect of $\rho$ on $r(t)$) for three values of $\rho$ over time ..........44
2.4 U.S. real market interest rates from 1930-2014 ..................................................51
2.5 U.S. annual consumption growth rate from 1930-2014 ..........................................51
2.6 Real interest rates, estimated $\rho(t)$ hat, 90 percent confidence intervals of $\rho(t)$ hat from 1930 to 2014 ..........................................................52
2.7 Estimate of the slope for U.S from 1930-2014 .....................................................54
3.1 The effects of increase in consumption growth uncertainty in certainty-equivalent $r(t)$ ......68
3.2 Certainty-equivalent $r(t)$ when $\rho$ hat is constant in discounting horizon (Type 1) ........70
3.3 Certainty-equivalent $r(t)$ when $\rho$ hat is non-constant in discounting horizon (Type 2) ....71
3.4 Schedule of certainty-equivalent $r(t)$ when $\rho$ hat=0.05 ........................................76
3.5 Schedule of certainty-equivalent $r(t)$ when $\rho$ hat=0.03 ........................................76
Chapter 1

Forming a Majority Coalition for Carbon Taxes Under a State-Contingent Pricing Mechanism

1.1 Introduction

The purpose of emission taxes is to restore efficiency when a negative externality is present. In the classical case, when damages are observable and can be traced to specific emitting activity, imposition of a Pigovian tax on the activity can result in a socially efficient outcome (Pigou, 1920; Baumol, 1972). But some externalities (such as global warming from greenhouse gases) have complex intertemporal features that make them much more difficult to deal with. The emissions do not directly affect welfare, instead they affect an environmental state variable, and the change in the state variable determines the magnitude of damages. Also, the emissions may affect the state variable only over a time lag of unknown length. These features give rise to two major difficulties for devising policy responses.

First, it can be difficult or even impossible to compute the optimal future tax rates. In the case of carbon dioxide and other greenhouse gases, complex prediction models are required that embed myriad assumptions about highly uncertain parameters (e.g. Nordhaus, 2007). The confidence intervals around key parameters are so wide as to yield an arbitrarily large range of marginal damage calculations. For instance, the IPCC (2007) Synthesis Report stated that peer-reviewed estimates of the social cost
of carbon ranged from $3 to $95 per ton of CO2 emissions, and other recent surveys show even wider ranges (e.g. Tol 2007; Golosov et al., 2014). These differences trace to divergent assumptions about climate sensitivity, response lags, discount factors etc. Though attempts to find tractable solutions to the problem continue (e.g. Golosov et al., 2014) they still require use of functional forms and parameter values that effectively assume away everything that makes the problem difficult in the first place (see Pindyck 2013 for a trenchant critique of such exercises). Incorporating Bayesian learning into a model may correct the initial parameter values over time. But the learning routine can take centuries to millennia to reach a 5% critical value, thus making it infeasible for climate policy (Kelly and Kolstad, 1999; Leach, 2007).

Second, the uncertainties that yield the wide range of conjectures about marginal damages imply individuals must have conflicting expectations of how the state variable will evolve over time. This in turn implies that they have divergent preferences over the optimal policy response. If the tax instrument must receive majority support in a voting system, and people can vote against it because it is too high or too low, forming a majority coalition can be effectively impossible.

The purpose of this paper is to examine how a recently-proposed solution to the first problem also addresses the second. McKittrick (2010, herein M10) suggested taking advantage of the information contained in observations of the state variable to calibrate a dynamic pricing rule that, under certain assumptions, can be expected to closely approximate the unobservable optimal tax path. In the M10 set-up, the environmental state variable at time \( t \) is a function of current and past emissions:

\[
s(t) = s(e_t, e_{t-1}, e_{t-2}, \ldots, e_{t-k})
\]

(1)
out to lag length \( k \), where \( k \) may be unknown. It is assumed that there are \( Q \) infinitesimally small emitters, which implies in any given period \( t \), total emissions are \( e_t = \sum_{j=1}^{Q} e_t^j \). In other words, equation (1) implies that the value of the state variable only depends on the emissions at time \( t \), \( t - 1 \) up to \( k \) lags. We do recognize, however, that the state variable may be a function of other variables, such as its own lagged values up to \( k \) lags, and other exogenous factors like assimilative ability of the environment to absorb emissions, the degree of heat absorption by the ocean which all may evolve over time. The main use of equation (1) is so that the model is tractable. This implies that the exact form of equation (1) may be different across countries and continents as exogenous factors noted above may vary by region to region.

Damages \( D \) are a function of the state variable, implying a valuation function of the externality in the form

\[
V(t) = \sum_{j=0}^{T} \beta^j D(s(t+j))
\]

(2)

where \( \beta \) is the discount factor and \( T \) is the policy planning horizon. The socially optimal price \( \tau(t) \) on emissions at time \( t \) is \( \tau(t) = \frac{\partial V}{\partial e_t} \), and a policy plan would consist of a sequence of future tax rates through to \( t+T \). Any attempt to derive the entire path at time \( t \) would run into the computational problems noted above, and any attempt to secure majority agreement to implement such a path would run into the difficulty that most observers would expect to get a price path higher or lower than the one they think is optimal. M10 derived a state-contingent pricing mechanism (herein SCPM) taking the form

\[
\bar{\tau}_t = \gamma s(t) \left( \frac{e_t}{\bar{e}_t} \right)
\]

(3)

where \( \bar{e}_t \) is a moving average of past emissions over the regulator’s best estimate of \( k \) periods and \( \gamma \) is a parameter that must be chosen to determine an initial value of the tax sequence. Using equation (2) (which describes the discounted present value of
and equation (1) (which describes evolution of state variable as function of emissions) and 6 simplifying assumptions, M10 derives the simple pricing rule for carbon emissions given by equation (3). M10 shows that over time, (3) is highly correlated to the unobservable optimal tax path implied by (1) and (2). To implement (3) the regulator only needs to know information available at time $t$, but instead of announcing a complete tax path, the policy implies a tax adjustment rule, so the future values of the tax will rise or fall in step with $s(t)$. Agents will therefore form expectations: those who expect $s$ to rise quickly (for instance under rapid global warming) will expect a rapidly increasing emissions price, whereas those who expect little change in $s$ will expect the emissions price to remain roughly unchanged from its initial value.

There remains the second problem, which is whether the SCPM instrument can obtain majority support for implementation. This paper focuses on that issue. We show that a conventional emissions tax based on political compromise across agents with divergent beliefs can easily fail to get majority support, but the SCPM typically does get majority support. We first examine a case in which voters are honest, in the sense that they only care about implementing the socially optimal tax rates, but they differ in their views of how the state variable will evolve over time. A voter who advocates for a high tax rate believes that current emissions will cause large changes in the state variable and therefore high marginal damages, while advocates for low taxes imply the opposite. Under this case, the SCPM always obtains 100 percent support and is robust to any distribution of beliefs. We then allow some voters to be dishonest, implying that they prefer either low or high tax rates irrespective of their beliefs about marginal damages, and we examine the incentives to support or reject the SCPM against an alternative in which the regulator implements a static compromise tax rate. We find that the SCPM may still obtain majority support, though it is not assured. However if all voters are dishonest, an unexpected outcome emerges in which, if the group that is
better off under the state-contingent option has a small majority then that option gets chosen, but as the size of the supporting majority increases, the vote eventually flips and the majority instead prefers a static option.

The remainder of the paper is organized as follows. Section 1.2 discusses earlier studies on voting mechanisms for public goods and externalities. Sections 1.3 and 1.4 develops the theoretic structure of our model and provides propositions and their proofs. Lastly, Section 1.5 presents conclusions.

1.2 Voting on taxes for externalities and public goods

Numerous authors have examined the way in which voting systems influence the adoption or rejection of proposed taxes. Pigouvian taxes aim to correct externalities and therefore can be welfare-improving. Revenue-raising taxes (e.g. Ramsey 1927) are imposed in such a way as to minimize the loss of private utility. The experimental results of Dresner et al (2006) show that the success of adopting a new tax policy depends on how well the voters understand the proposed policy. For example, some voters may not support a Pigouvian tax because they do not fully understand how it is used to enhance efficiency. Similarly, Clinch et al (2006) conclude that public trust in the government plays a key role in determining the support for new taxes. Several natural field experiments have shown that framing affects voting behavior. For instance, according to McCaffery and Baron (2003), some people may react negatively even to the use of the word “tax”. On the other hand, Sausgruber and Tyran (2005) showed experimentally that some people prefer indirect over direct taxes, which they call “fiscal illusion.”

An earlier, related literature examined positive externalities such as publicly funded education. Creedy and Francois (1990) showed that if education provides a positive externality to the economy by inducing economic growth, and if only certain (high)
type of individuals can benefit from education, then under certain conditions a majority of uneducated individuals would be willing to pay taxes to subsidize education for high types in return for (higher) economic growth. Johnson (1984) draws the same conclusion, however, his model does not incorporate opportunity cost of education in terms of forgone wage.

Alesina and Passarelli (2013) analyze majority voting outcomes when the government has three environmental policy tools: a rule, a quota and a proportional tax. In their paper, they define a rule as an instrument that sets an upper limit to the activity and a quota as that which requires a proportional reduction of the activity. They show that majority voting may not yield a socially optimal outcome when there are several policy options and voters have divergent preferences. For example, when the group producing the externality is relatively small, the majority supports a rule rather than the tax that would be chosen by the social planner. On the other hand, if this group is relatively large, then the majority supports a tax whereas the social planner would choose a rule. In other words, if the group responsible for the externality is in the minority, then the majority will choose a policy that puts the greatest burden of reducing the externality to the minority group, and vice versa. These results are in line with those of Schneider and Volkert (1999) who show that when the voting community is composed of groups with differentiated interests, then the voting outcome may not be socially optimal.

Fredriksson and Sterner (2005) incorporate differences in abatement technologies across firms and show that “clear” firms may lobby for higher tax rates if the revenue is used for rebates. Kawahara (2011) builds a model with assumptions that voters do not observe politician types and environmental damage. Under his model, a pooling equilibrium results in a sub-optimal tax rate, whereas in the separating equilibrium,
pro-environmental politicians choose a tax rate that is too high in order to distinguish themselves from other types. Lastly, Cremer, De Donder and Gahvari (2004) examine revenue recycling and voting outcomes. They show that if environmental tax revenue is used to subsidize income and capital taxes, then the majority will choose an environmental tax that is too low.

Overall, the literature points to intuitively plausible results that, whether the majority are voting on a Pigovian tax to enhance efficiency or on income taxes to raise revenue for publicly-funded education, the voting outcome depends not only on preferences of the voters but also on the perceived distribution of benefits (and/or costs). The first framework we examine below provides a strikingly different outcome, in which a state-contingent emissions tax obtains 100 percent support regardless of the distribution of beliefs about the benefits. Current models are also characterized by yielding monotonic effects, in the sense that the larger the majority in favour of an option, the more likely it obtains voter approval. Our second case allows for all voters to be dishonest, and there we find an interesting non-monotonicity in the results. Conditions that yield majority support for an environmental tax require that the majority not be overly large: if it grows past a certain point the vote flips.

1.3 The voting model

1.3.1 Voting on the simple static case

We now examine the question of whether self-interested voters would support a proposed tax policy. The voting environment is described as follows. There are $N$ voters indexed by $i \in N = \{1, \ldots, N\}$ and one policy maker who proposes a tax rate $q$. Each voter chooses an action $a_i \in \{1,0\}$, denoting a vote in favour ($a_i=1$) or against ($a_i=0$), to minimize a quadratic loss (utility) function. The proposed tax rate is implemented
if it obtains majority support.

1.3.2 A simple static case

A voting community is made up of $N$ persons, labeled $i = \{1, \ldots, N\}$ where $N>2$. Each voter has a single-peaked preference over the set of tax rates, which we take to be the real line $\mathbb{R}_+$. Their decision depends on the distance between the proposed tax rate and their ideal tax rate $\tau_i$ as well as cut-off value $d$. We denote the loss of individual $i$ when the tax rate is $q \in \mathbb{R}_+$ by $L_i(q, \tau_i) = (\tau_i - q)^2$. Voters choose their action to minimize this value.

A voter with an ideal point $\tau_i$ will choose to support the proposed tax rate if and only if the loss is less than or equal to a cut-off value $d$. We can also interpret this as value of reservation disutility when proposed tax rate does not pass.\footnote{Due to quadratic nature of the loss function, disutility of proposed tax rate when $\tau_i \neq q$ is symmetric around $\tau_i$.} Therefore, a voter’s decision rule, assuming $d$ constant, is as follows:

$$
\begin{align*}
    a_i &= 1 \quad \text{if} \quad L_i \leq d \\
    a_i &= 0 \quad \text{if} \quad L_i > d
\end{align*}
$$

The above condition implies that the following must hold for a voter with preferred tax rate $\tau_i$ who chooses $a_i=1$,

$$(\tau_i - t)^2 \leq d$$

$$\Rightarrow \tau_i \in [t - \sqrt{d}, t + \sqrt{d}]$$
This result shows, in particular, that the proposed tax rate can not be too far from voter $i$’s most preferred rate, and a higher value of $d$ increases the propensity to vote yes.\footnote{Note that this result is different from voting models based on a median voter’s prospective. In this case, all voters with preferences located left of a median voter would support the proposed tax, and those with preferences located right side of a median voter would reject it. As a result, our results are strikingly different from median voter framework.}

1.3.3 The policy maker’s problem

The policy maker’s objective is to choose the tax rate $q$ to minimize the sum of losses of voters.

\[
\mathcal{L}(q) = \arg \min_q \sum_{i=1}^{N} L_i \\
= \arg \min_q \sum_{i=1}^{N} (\tau_i - q)^2
\]

Differentiating the objective function with respect to $q$, we obtain the first order condition.

\[
\frac{\partial \mathcal{L}}{\partial q} = -\sum_{i=1}^{N} 2 \ast (\tau_i - \hat{q}) = 0
\]

\[
\hat{q} = \frac{1}{N} \sum_{i=1}^{N} \tau_i
\]

\[
\hat{q} = \bar{\tau} \tag{4}
\]

Equation (4) shows that the policy maker should select the mean of preferred rates. Total losses can be derived by combining equation (4) in the objective function:

\[
\mathcal{L} = \sum_{i=1}^{N} (\tau_i - \bar{\tau})^2
\]
and rearranging,
\[ \mathcal{L} = (N - 1) * \sigma_\tau^2 \]  
(5)

where \( \sigma_\tau^2 \) is variance of preferred tax rates among all voters. Differentiating the loss function with respect to \( N \) and \( \sigma^2 \), we obtain following conditions:

\[
\frac{\partial \mathcal{L}}{\partial \sigma^2} = N - 1 > 0 \\
\frac{\partial \mathcal{L}}{\partial N} = \sigma^2 > 0
\]  
(6)

Condition (6) says that total losses are strictly monotonic in \( N \) and variance of \( \tau \).

1.3.4 The relationship between the variance of \( \tau_i \) and voting outcomes

It will be convenient to define voter \( i \)'s propensity to vote against \( q \) as:

\[ p_i = \frac{|\tau_i - \bar{\tau}|}{\sqrt{d}} \]  
(7)

where

\[ 0 < p_i \leq 1 \quad \Rightarrow \quad a_i = 1 \quad (\text{low propensity to vote against}) \]

\[ p_i > 1 \quad \Rightarrow \quad a_i = 0 \quad (\text{high propensity to vote against}) \]

Let \( \varepsilon_i \) be defined as follows:

\[ \varepsilon_i = |\tau_i - \bar{\tau}| \]

Combining this with equation (7), the variance of \( \tau \) and the model’s parameters \( (N, \)
\( d \) can be written

\[
\sigma^2_\tau = \frac{d}{(N-1)} \sum_{i=1}^{N} p_i^2
\]

For a fixed values of \( d \) and \( N \), an increase in \( \sigma^2_\tau \) requires an increase in at least one \( p_i \). In other words, the higher the variance, the higher the propensity for at least one person to vote no.

Let us consider following four possible cases:

- case 1) There is unanimous yes votes (\( a_i=1 \) \( \forall \) \( i \in N \)).
- case 2) There is unanimous no votes (\( a_i=0 \) \( \forall \) \( i \in N \)).
- case 3) There is majority yes votes.
- case 4) There is majority no votes.

**Case 1** In order to observe unanimous yes votes, the following condition must be satisfied:

\[
\begin{align*}
    &p_i \leq 1 \quad \forall i \\
    \Leftrightarrow & \varepsilon_i^2 \leq d \quad \forall i \\
    \Leftrightarrow & \frac{1}{N-1} \sum_{i=1}^{N} \varepsilon_i^2 \leq \frac{N*d}{N-1} \Leftrightarrow d \quad (i) \\
    \Leftrightarrow & \sigma^2_\tau \leq d \quad \text{for large } N \quad (ii)
\end{align*}
\]

Therefore, observing an unanimous yes vote implies that the variance of \( \tau \) is less than \( d \) for all voters.

However, \( \sigma^2_\tau \leq d \) does not necessary implies unanimous yes votes.

Proof by contradiction: Suppose \( N=5 \) where \( \tau=(1,1,2,2,4) \) is a vector of preferred
tax rates for voters 1 though 5 respectively, and let \( d = 3 \). The following shows that condition (i) is satisfied.

\[
\frac{1}{N-1} \sum_{i=1}^{5} \epsilon_i^2 = 1.5 \leq \frac{N \cdot d}{N-1} = 3.75
\]

However \( p_5 = \frac{|4-2|}{\sqrt{3}} = 1.15 > 1 \). Thus voter 5 rejects the proposal. QED.

**Case 2**) Similarly, observing unanimous no votes implies the following:

\[
\begin{align*}
\{ & p_i > 1 \; \forall i \\
\Rightarrow & \epsilon_i^2 > d \; \forall i \\
\Rightarrow & \frac{1}{N-1} \sum_{i=1}^{N} \epsilon_i^2 > \frac{N \cdot d}{N-1} \approx d (iii) \\
\Rightarrow & \sigma^2 > d \; for \; large \; N (iv)
\end{align*}
\]

However, \( \sigma^2 > d \) does not necessary implies unanimous no votes.

Proof by contradiction: Suppose \( N = 5 \) where \( \tau = (1,1,7,7,4) \) is a vector of preferred tax rates for voter 1 though 5 respectively, and let \( d = 3 \). The following shows that condition (iii) is satisfied.

\[
\frac{1}{N-1} \sum_{i=1}^{5} \epsilon_i^2 = 9 > \frac{N \cdot d}{N-1} = 3.75
\]

However \( p_5 = \frac{|4-4|}{\sqrt{3}} = 0 < 1 \). Thus voter 5 supports the proposal. QED.
Cases 3) and 4) Let us now consider a case in which we observe both yes \((a_i=1)\) and no \((a_i=0)\) votes. Let \(M\) and \(R\) denote the total number of yes votes and no votes respectively. Arranging voters in terms of \(\varepsilon\) in a monotonically increasing manner, we summarize the voters as the following:

\[
\sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{M} \varepsilon_i^2 + \sum_{j=M+1}^{N} \varepsilon_j^2
\]

The above equation can be written as

\[
\sigma_r^2 = \frac{1}{N-1} \frac{M-1}{M-1} \sum_{i=1}^{M} \varepsilon_i^2 + \frac{1}{N-1} \frac{R-1}{R-1} \sum_{j=M+1}^{N} \varepsilon_j^2
\]

\[
= \frac{M-1}{N-1} \sigma_y^2 + \frac{R-1}{N-1} \sigma_n^2
\]

(8)

where \(\sigma_y^2\) and \(\sigma_n^2\) are the variances of yes and no voters respectively.

Finally, for large values of \(M\) and \(R\), we obtain the following condition:

\[
\sigma_r^2 \approx \frac{M}{N} \sigma_y^2 + \frac{R}{N} \sigma_n^2
\]

(9)

Case 3) Suppose we observe majority yes votes: \(M>R\), which implies \(\frac{R}{M}<1\).

Rearranging equation (9), we obtain the following condition:

\[
\sigma_y^2 = \frac{N}{M} \sigma_r^2 - \frac{R}{M} \sigma_n^2
\]

\[
= \frac{M+R}{M} \sigma_r^2 - \frac{R}{M} \sigma_n^2
\]

\[
= \sigma_r^2 + \frac{R}{M} (\sigma_r^2 - \sigma_n^2)
\]

(10)

\[
\sigma_y^2 = \sigma_r^2 - \frac{R}{M} (\sigma_n^2 - \sigma_r^2)
\]
Finally,
\[
\sigma^2_y < 2\sigma^2_\tau - \sigma^2_n \tag{11}
\]
since \( \frac{R}{M} < 1 \). Conversely, observing condition (11) implies we have majority yes votes. This is because condition (11) is only true when we impose the condition \( M > R \). Instead if we impose \( R \geq M \), equation (10) becomes\(^3\)
\[
\sigma^2_y \geq 2\sigma^2_\tau - \sigma^2_n
\]
Therefore, observing condition (11) is a necessary and sufficient condition for observing majority yes votes. Condition (11) can also be written as the following:
\[
\sigma^2_\tau > \frac{1}{2}(\sigma^2_y + \sigma^2_n)
\]
In other words, the variance of preferred tax rates \( \tau \) must be greater than the unweighted average variances of preferred tax rates among yes and no votes.

\textit{Case 4)} Observing majority no votes implies that \( M < R \). Therefore, imposing this condition on equation (9), we obtain the following condition:
\[
\sigma^2_y > 2\sigma^2_\tau - \sigma^2_n \tag{12}
\]
Observing condition (12) implies we have majority no votes since this condition is only true when we impose the condition that \( M < R \). Therefore, observing condition (12) is necessary and sufficient condition for observing majority no votes. In other words, the variance of preferred tax rates \( \tau \) must be less than the unweighted average variances of preferred tax rates among yes and no votes.\(^3\)

\(^3\)There is a tie if \( R = M \) and majority no votes if \( R > M \)
As shown in equations (11) and (12), the key factors are the variances of the preferred tax rates among yes, no and all voters. However, it is not clear how marginal increase or marginal decrease in the variances of the preferred tax rates among yes and no votes affect the majority outcome. This is because these variances also affect the overall variance of preferred tax rates. Nonetheless, as the variance of overall preferred tax rates increases, we must converge to case 2 where we observe unanimous no vote. This implies that it will be hard for a static tax policy to obtain majority support if the variance of preferred tax rates is large.

1.4 Voting outcomes when some of voters are honest and others are not

1.4.1 The current problem with static CO$_2$ emission tax policy

Efforts to address the complex intertemporal externality of global warming has resulted in failure to obtain majority support for a carbon tax. Anecdotally, this is in part due to the wide variance of preferred tax rates, as discussed in the introduction, leading to a high propensity to vote against any specific proposal. The variance of preferred tax rates arises because CO$_2$ emissions do not cause direct harm, but have an indirect effect through an uncertain influence on the state variable. The problem appears to be intractable, unless preferences over preferred tax rates converge. In the next section we show that an alternative solution arises using a different policy proposal.

1.4.2 The loss function under the SCPM

The loss function of voter $i$ under the SCPM has the same form as before, but $\tau$ is replaced with the announced tax rule: $q_t = f(s_t),$

$$L_i^{sc} = (\tau_i - f(s_t))^2$$ (13)
where \( \tau_i \) is the private preferred tax rate for voter \( i \) and \( f \) is given by equation (3). Since the SCPM is a rule based on an observable state-variable, individuals will form expectations of the state-variable in order to form expectations of the path of tax rates.

We assume that individual voters have an exogenously-given private information set (denoted \( I_i \)) that governs their forecast of the state-variable. A voter’s information set is assumed to yield private beliefs on the functional form of \( s(t) \) and the sensitivity of the state-variable to current and past emissions \( \left( \frac{\partial s}{\partial e_{t-j}} \right) \). We assume that all voters have the same beliefs about the form of the damage function. This implies that, given an observed value of the state variable at time \( t \), everyone would agree on what the optimal tax rate should be. For instance, if we observe in year 2050 that the value of the state variable has significantly increased, then everyone would agree that the tax rate should have increased too, and vice versa. Nonetheless, since voters may have different beliefs about the relationship between the state variable and emissions as well as the unknown lag length \( (k) \), it is possible for voters to project different optimal tax rates.

1.4.3 Voting outcome when all voters are honest

We now examine cases in which the regulator proposes the SCPM instead of a static tax rate. We first examine a case in which voters are honest. The results are strikingly different than in previous studies, in the sense that the SCPM obtains 100% support regardless of the distribution of beliefs about the benefits. We then allow some of the voters to be dishonest and examine conditions in which the SCPM still obtains majority support. Lastly, we analyze a case in which all voters are dishonest. We also observe a non-conventional outcome here, where conditions that yield majority support for the SCPM require that the majority not be overly large.

Proposition 1). Suppose all voters \( i \in N \) have the same beliefs regarding the dam-
age function but have different beliefs about how the state variable will evolve over
time in response to emission $e_t$. Also suppose voters are honest, so each one prefers the
socially-optimal value implied by their views on the state variable. If agents have unbi-
ased forecasts of $s$, the SCPM will obtain 100% support regardless of the distribution
of beliefs about the effect of emissions on $s_t$.

Proof of Proposition 1).

The preferred tax $\tau_i$ is based on $E_i(s/I_i)$ where $s$ is a vector of future realizations
of the state variable and $I_i$ is the private information set for voter $i$. Here, we as-
sume that individuals have different expectations about $s$ because of exogenously given
information set that is unique to each individual ($I_i$). Specifically,

$$\tau_i \equiv E_i(V'(s)) = V'(E_i(s/I_i))$$

where $V'$ is $\frac{\partial V}{\partial e_t}$ as defined in the preliminary section. The regulator proposes to follow

$$q_t = f(s_t)$$

where $f$ is chosen to approximate $V'$ using equation (3).

Then at time $t$, the following is true:

$$E(L_i) = E_i((q_t - \tau_t)^2)$$

$$= E_i(q_t^2 - 2q_t\tau_t + \tau_t^2)$$

$$= E_i(q_t^2) - 2E_i(q_t\tau_t) + E_i(\tau_t^2)$$

$$= E_i(V'(s)^2) - 2E_i(V'(s))V'(E_i(s/I_i)) + E_i(V'(E_i(s/I_i))^2).$$

(14)
Since
\[ E_i(V'(s)) = V'(E_i(s/I_i)) \]
then
\[ E(L_i) = 0 \]
and
\[ a_i = 1 \forall i \in N \quad (15) \]

QED.

Condition (15) arises because everyone expects their preferred tax rate to be implemented. This is a result of the SCPM being able to provide a close approximation of the socially optimal tax rate, so if all voters want to implement the socially optimal tax rate at every time \( t \), then it does not matter whether voter \( i \) believes that the state variable will increase or decrease in the future. Those who believe in a rapid increase in the \( s(t) \) would anticipate a steep increase in the tax rates while those who believe in little to no increase in \( s(t) \) expect low tax rates. In the end, under the SCPM all voters expect the actual tax path to be their preferred one.

1.4.4 Voting outcomes when some or all voters are dishonest

The following Propositions (2)-(5) involve voting outcomes where some of the voters have preferred tax rates independent of the socially optimal rate (i.e. they prefer either a zero tax or a maximum tax on a priori grounds, irrespective of marginal damages). We assume that there are three types of policy preferences among voters: Type 1 prefers low values for \( \tau \) regardless of \( s \); Type 2 prefers the socially optimal rate, and Type 3 prefers high values for \( \tau \) regardless of \( s \). Therefore we call types 1 and 3 dishonest voters. Next we assume that all possible tax rates are bounded between 0 and some finite number \( w \): \( 0 \leq \tau \leq w \). Thus, the privately optimal tax rates are: \( \tau_1^* = 0 \), \( \tau_2^* = V' \) and
\( \tau_3^* = w \) for types 1, 2 and 3 respectively. Denote the numbers in each group as \( n_1, n_2 \) and \( n_3 \).

Suppose the regulator proposes the following two rules.

Rule 1 (Static option): Minimize the loss function using \( q=\overline{\tau} \). Given this rule, the following summarizes the aggregate losses for each type.

\[
\begin{align*}
L_1^1 &= n_1 * (\overline{\tau} - 0)^2 \\
L_2^1 &= n_2 * (\overline{\tau} - V')^2 \\
L_3^1 &= n_3 * (\overline{\tau} - w)^2
\end{align*}
\]

The total loss under Rule 1 is

\[
L^1 = n_1 \overline{\tau}^2 + n_2 * (\overline{\tau} - V')^2 + n_3 * (\overline{\tau} - w)^2
\]

Rule 2: The state-contingent approach, where \( q=V' \). As above, the following summarizes aggregate losses for each group.

\[
\begin{align*}
L_1^2 &= n_1 V'^2 \\
L_2^2 &= n_2 (V' - V')^2 = 0 \\
L_3^2 &= n_3 (V' - w)^2
\end{align*}
\]
The total loss under Rule 2 is

\[ L^2 = n_1 V''^2 + n_3 (V' - w)^2 \]

The reduction in losses (\( \Delta_i \)) by going from \( q = \bar{\tau} \) to \( q = V' \) are as follows:

For group \( n_1 \),

\[ \Delta_1 = L^1_1 - L^2_1 = n_1 (\bar{\tau}^2 - V''^2) \]

For group \( n_2 \),

\[ \Delta_2 = L^1_2 - L^1_2 = n_2 (\bar{\tau} - V')^2 \]

For group \( n_3 \),

\[ \Delta_3 = L^1_3 - L^2_3 = n_3 (\bar{\tau} - w)^2 - n_3 (V' - w)^2 \]

\[ = n_3 [(\bar{\tau} - w)^2 - (V' - w)^2] \]

\[ = n_3 [\bar{\tau}^2 - 2w\bar{\tau} + w^2 - V''^2 + 2wV' - w^2] \]

\[ \Delta_3 = n_3 [(\bar{\tau}^2 - V''^2 + 2w(V' - \bar{\tau})] \]

Lemma 1) \( V' > \bar{\tau} \) implies that \( V' > \frac{n_3 w}{n_1 + n_3} \) and vice versa.

Proof.

Since \( \bar{\tau} \) is the average of preferred tax rates of \( N = \{n_1, n_2, n_3\} \), it is equal to the
following,

\[
\bar{\tau} = \frac{n_1(0) + n_2 V' + n_3 w}{n_1 + n_2 + n_3} = \frac{n_2 V' + n_3 w}{n_1 + n_2 + n_3}
\]

It follows that when \( V' > \bar{\tau} \),

\[ V' > \frac{n_2 V' + n_3 w}{n_1 + n_2 + n_3} \]

Solving for \( V' \), we obtain the following condition:

\[ V' > \frac{n_3 w}{n_1 + n_3} \]

Similarly, when \( V' < \bar{\tau} \) we obtain that

\[ V' < \frac{n_3 w}{n_1 + n_3} \]

QED.

For what follows it would be helpful to visualize the tax rates as points on a line, as in Figure 1.1.

![Fig 1.1. Case where \( V' > \bar{\tau} \).](image)

Proposition 2). If \( V' > \bar{\tau} \), the SCPM obtains majority support if \( n_2 + n_3 > n_1 \).

Proof.
The decrease in losses ($\Delta_i$) by going from $q=\bar{\tau}$ to $q=V'$ are as follows:

For group $n_1$,

$$\Delta_1 = n_1(\bar{\tau}^2 - V'^2) < 0 \quad (16)$$

For group $n_2$,

$$\Delta_2 = n_2(\bar{\tau} - V')^2 > 0 \quad (17)$$

For group $n_3$,

$$\Delta_3 = n_3[\bar{\tau}^2 - V'^2 + 2w(V' - \bar{\tau})]$$

$$= n_3[(\bar{\tau} + V')(\bar{\tau} - V') - 2w(\bar{\tau} - V')]$$

$$= n_3(\bar{\tau} - V')(\bar{\tau} + V' - 2w)$$

Firstly, by assumption $\bar{\tau} - V' < 0$. Secondly, since $\bar{\tau} < V' < w$, we obtain the condition $\bar{\tau} + V' - 2w < 0$. Therefore $\Delta_3 > 0$.

Conditions (16), (17) and (18) show that groups 2 and 3 vote for the SCPM while group 1 does not. Thus in order to obtain majority support, the following condition must hold true:

$$n_2 + n_3 > n_1$$

QED.

Proposition 3). Given $V' < \bar{\tau}$, the SCPM obtains a majority support if $n_1 + n_2 > n_3$. 
Proof.

Following the set up as in Proposition 3), the decrease in losses by group are as follows:

For group $n_1$, 

$$\Delta_1 = n_1(\bar{\tau}^2 - V'^2) > 0$$ (19)

For group $n_2$, 

$$\Delta_2 = n_2(\bar{\tau} - V')^2 > 0$$ (20)

For group $n_3$, 

$$\Delta_3 = n_3[\bar{\tau}^2 - V'^2 + 2w(V' - \bar{\tau})]$$ (21)

$$= n_3[(\bar{\tau} + V')(\bar{\tau} - V') - 2w(\bar{\tau} - V')]$$

$$= n_3(\bar{\tau} - V')(\bar{\tau} + V' - 2w)$$

First, by assumption $\bar{\tau} - V' > 0$. Second, since $V' < \bar{\tau} < w$, we obtain that $\bar{\tau} + V' - 2w < 0$. Therefore $\Delta_3 < 0$. Conditions (19), (20) and (21) show that to obtain majority support for the SCPM, following condition must hold true:

$$n_1 + n_2 > n_3$$

QED.
Proposition 4). If $V=\bar{\tau}$ then voters are indifferent between the two policies.

Proof.

It follows that when $V'=\bar{\tau}$, conditions (19), (20) and (21) become $\Delta_1=\Delta_2=\Delta_3=0$.

QED.

Proposition 5). Suppose all voters are dishonest. If the group that is better off under the state-contingent option has a small majority then that option gets chosen, but as its majority increases the vote eventually flips to prefer the static option.

Proof.

When $n_2=0$ the average of preferred tax rates is as follows:

$$\bar{\tau} = \frac{n_3w}{n_1+n_3} \tag{22}$$

and the limit properties with respect to $n_1$ and $n_3$ are as follows:

$$as \ n_1 \to \infty, \ \bar{\tau} \to 0$$
$$as \ n_1 \to 0, \ \bar{\tau} \to w$$

$$as \ n_3 \to \infty, \ \bar{\tau} \to w$$
$$as \ n_3 \to 0, \ \bar{\tau} \to 0$$

Suppose initially $n_1>n_3$ and $\bar{\tau}>V'>0$. By Proposition 3) the SCPM obtains majority support. But let $n_1$ get larger. It follows by condition (22) that as $n_1$ increases, we must converge to $0<\bar{\tau}<V'$ and $n_1>n_3$. By Proposition 2) the SCPM fails to obtain
Conversely, suppose initially $n_3 > n_1$ and $V > \bar{\tau} > 0$. By Proposition 2) the SCPM obtains majority support. However, condition (22) implies that as $n_3$ increases, we must converge to $\bar{\tau} > V > 0$ and $n_3 > n_1$. By Proposition 3) the SCPM fails to obtain majority support.

QED.

These results suggest that even when all voters have preferred tax rates independent of the socially optimal rate, the SCPM may still be able to obtain majority support and implement the socially optimal tax rates for all $t$. If the group that is better off under the state-contingent option has a small majority then that option gets chosen, but as its majority increases the vote eventually flips to prefer the static option.

1.5 Conclusions

This paper has examined voting outcomes when the policy maker proposes an externality pricing instrument that is either based on a static political compromise or on a state-contingent updating rule. Firstly, in the case where all voters are honest (voters prefer the socially optimal price based on their expectation of marginal damages), a standard pricing mechanism aimed at minimizing political losses may never obtain majority support, depending on the variance of private expectations about the level of marginal damages. On the other hand, the SCPM always obtains majority support regardless of voters’ beliefs about the future evolution of the state variable and its marginal damages. The SCPM is able to attain this outcome because everyone expects their private optimal tax to be implemented at each point in time. As a result, the degree of polarization in voters’ beliefs is no longer an obstacle for majority support.
Secondly when we allow some of the voters to be dishonest (they prefer either a zero tax or a maximum tax on a priori grounds, irrespective of marginal damages), we find that under certain conditions the SCPM obtains majority support. However if all voters are dishonest \((n_2=0)\), an unexpected outcome emerges in which, if the group that is better off under the SCPM option has a small majority then that option gets chosen, but as its majority increases the vote eventually flips to prefer the static option.
Chapter 2

Discounting when The Consumption Path Is Affected by the Rate of Time Preference

2.1 Introduction

In the last few decades, there has been an enormous growth in both theoretical and empirical literature on the social discount rates. This is attributed to the fact that economists are increasingly being asked about the choice of the social discount rate to be used in Cost-Benefit Analysis (herein CBA). The first formal theoretical derivation of the social discount rate was published in 1928 by a notable economist Frank Ramsey. The work is based on neoclassical growth model framework without uncertainty and has become the benchmark case in the discounting literature. Since then, numerous theoretical papers that extend Ramsey growth framework have emerged (e.g Gollier, 2002a, 2002b, 2008; Mankiw 1981; Pindyck and Wang, 2012; Weitzman, 2012).

The emergence of extremely long-lived policies and projects such as carbon tax and the construction of nuclear waste disposal sites to meet the energy demand mean that assessment of CBA could hinge on the social discount rate used. Relatively small differences in this rate may result in an enormous difference in the policy assessment. Not surprisingly, the discount rate issue has become a major source of disagreement and is one of the most critical issues in economics. For example, Stern (2006) published
a report for the British government which calls for an immediate implementation of climate change mitigation activities that reduce greenhouse gas emissions by about 3% per year. The report suggests devastating consequences if no action is taken now that are equivalent to losing at least 5 percent and up to 25 percent of global GDP for all years from now on. On the other hand, critics of the report argue that Stern’s results are due to an extremely low discount rate (1.3%) employed in the analysis and that his recommendations are no longer supported with a higher social discount rate (see, Gollier et al. 2008; Mendelsohn, 2008; Nordhaus, 2007b). Guo et al. (2006) have found that the discounting rate is responsible for a difference of a factor of 40 in the estimated social cost of carbon. Even greater differences have been observed in other studies (see Tol (2005) for a good survey of past estimates).

The disagreements about the social discount rate are not merely about how to empirically estimate the terms in Ramsey equation, but there are major disagreements stemming from ideological differences as well. The latter issue is mainly from the fact that the rate of time preference is not an observable variable in practice. The Ramsey equation is as follows,

$$r = \rho + \eta \frac{\dot{C}}{C}$$

where $r$ is the social discount rate, $\rho$ is the pure rate of time preference, $\eta$ is the elasticity of marginal utility of consumption and $\frac{\dot{C}}{C}$ is the consumption growth rate. There is an important distinction between the social discount rate, also known as consumption discount rate, and $\rho$, also called utility discount rate. $\rho$ is a parameter reflecting the level of impatience or psychological preference for the presence of a representative agent, whereas the social discount rate is the rate used to discount future consumption by a social planner.

The third chapter of this thesis examines the effects of uncertainty in $\rho$ on the social discount rates in the presence of consumption uncertainty.
The main purpose of this paper is to examine the relationship between $\rho$ and the social discount rate in a general equilibrium framework. In the literature, the consumption growth rate is commonly treated to be fixed in Ramsey equation, which implies unit slope between $\rho$ and $r$. We relax this assumption and analyze their relationship by examining the slope, $\frac{\partial r}{\partial \rho}$. Our objective is important for two reasons. First, since it is the social discount rate that matters in CBA, understanding the relationship between $\rho$ and the social discount rate is vital and critical to the informed choice of $\rho$ as well as the debates about the issues noted earlier. As an illustrative example, suppose $r$ is invariant to changes in $\rho$. This relationship implies that choice of $\rho$ would not matter as it does not influence the value of $r$ and debates about $\rho$ stemming from philosophical or ideological differences would settle immediately. Second, by carefully examining $\frac{\partial r}{\partial \rho}$ especially for the closed form solution case, we can be better informed about the factors that influence this relationship which may be invaluable information to policy makers.

In the general model, we find that the sign of $\frac{\partial r}{\partial \rho}$ is ambiguous, and it is only equal to one under certain conditions. Since we can not obtain general results in the previous set up (without being able to derive a functional form of $\frac{\partial r}{\partial \rho}$), we resort to a closed form model where we observe that the shape of $\frac{\partial r}{\partial \rho}$ depends on several exogenous parameters and their levels.

Finally, we provide some empirical evidence on this issue. In this section, we apply the Time-Varying Semi-Parametric Smooth Coefficient model to first estimate the parameters in Ramsey equation, namely $\rho$ and $\eta$. We use data on long-term, high-quality government and corporate bonds and the consumption growth rate data for the U.S from years 1930 to 2014. Using the estimated parameter values, we propose a simple way to estimate $\frac{\partial r}{\partial \rho}$ and report its results.
The paper is organized as follows. Section 2.2 is the literature review. In Section 2.3, we present the general model framework, followed by a special closed form model. Section 4 provides empirical results for the U.S. Section 2.5 is the conclusion.

2.2 Past theoretical and empirical works on the social discount rates

Since the publication of Frank Ramsey’s “A Mathematical Theory of Savings” in 1928, the literature on the social discount rate has grown enormously. Especially in the last few decades, with the emergence of issues like global warming where the effects of a policy could be felt for centuries and millennia, the role of the discount rate in CBA has become ever more critical and vital. Numerous authors have relaxed some of the assumptions embedded in the original Ramsey framework (e.g. Baumgartner et al., 2015; Guo et al., 2006; Hoel and Sterner 2007; Gollier 2002a,2002b,2008; Weikard and Zhu, 2005; Weitzman, 2012).

In Ramsey framework, there are two main determinants of the level of the social discount rate. One is the wealth effect, which is based on the idea that under the usual assumption of decreasing marginal utility of consumption, an additional unit of consumption is less valuable (generates less utility) to the future generations (than the current one) if the consumption growth is positive over time. This is the main argument for placing a higher discount rate for the future. The other determinant is the precautionary effect, which lowers the social discount rate for the future when there is uncertainty about the future and people are prudent. For instance, when the growth

5Prudence refers to the propensity to accumulate savings in the face of future uncertainty. An agent is prudent if the third derivative of his utility function is positive.
rate of consumption is assumed to be stochastic and follows a certain process over the
discount horizon (as opposed to being deterministic), the precautionary effect reduces
the social discount rates. (e.g. Mankiw, 1981; Gollier 2002a, 2002b, 2008; Weitzman,
2012). Gollier (2002b) has shown that if the preference function exhibits decreasing
relative risk aversion or decreasing absolute risk aversion, then uncertainty in the con-
sumption growth rate reduces the social discount rate at any time horizon. He also
shows that the discount rate is decreasing for more distant futures. However, some
empirical studies have shown that this effect is likely to be small for the U.S (e.g.
Kocherlakota, 1996; Gollier, 2008).

Guo et al. (2006) examine the consumption uncertainty in a discrete framework. They
accomplish this by assuming that for any given year there are three possible consump-
tion growth rates, and then by calculating corresponding certainty-equivalent discount
rates. An alternative scheme is also considered where the consumption growth rates
follow a normal distribution so that there are five possible growth rates in each year.
The results show higher consumption uncertainty decreases the certainty-equivalent
discount rate under certain conditions.

A more recent, related literature examines the effects of environmental quality on so-
cial discount rates (Baumgartner et al., 2015; Hoel and Sterner 2007; Weikard and Zhu,

\textsuperscript{6}Note that the usage of the time argument may be different depending on the context. When we refer to discounting time horizon, or over time horizon, this means comparing social projects or policies with different maturities. We can also interpret this as the remaining life of the project as we move in time. For example, suppose a project is expected to last 50 years. At time zero, one could calculate the optimal amount of investment for each period (and corresponding consumption levels) for the next 50 years. However after one year, the time horizon has changed to 49 years (not 50 years), so we would be dealing with a shorter discounting horizon. The usual result in the literature is that a non-constant discount rate results in the time inconsistency problem. If the time argument is defined as the time index, referred to as ’in time’, it means comparing the marginal discount rates between two adjacent periods. For example, if discount rate is decreasing in time, then this would imply that the marginal social discount rate between year 1 and 2 is larger than its rate between year 50 and 51.
Weikard and Zhu have shown if manufactured and environmental goods are not substitutable, then an accounting price cannot exist in which each good will have its unique discount rate, a phenomenon known as 'dual discounting.' In general decline in environmental quality is shown to lower the social discount rate. Based on the theoretical results of Weikard and Zhu (2005), Baumgartner et al. (2015) have estimated the difference in the social discount rates between manufactured consumption goods and ecosystem services. Using panel data for five countries (Brazil, Germany, India, Namibia, UK), they showed that the estimated discount rates for manufactured goods are consistently higher than for ecosystem services for all countries, the greatest gap in the two rates being observed in less developed countries. The results seem to suggest a convergence of the two rates as income level goes up.

Pindyck and Wang (2012) have incorporated possibilities of catastrophic events, such as wars and other epidemics into the model. Their framework is similar to that incorporating the consumption uncertainty, in the sense that increasing possible (extreme) outcomes serves to increase its uncertainty, thereby further reducing the discount rate. The catastrophic risk is modeled as a Poisson process with a particular arrival rate of the 'bad' events, and when it occurs, the consumption falls by random percentage points. In a related study, Gollier et al. (2008) show that if shocks to the consumption are positively correlated, the social discount rate will decline over discount horizon if the preference exhibits constant relative risk aversion (CRRA). However, if the uncertainty in the growth rate of consumption is assumed to be independent and identically distributed with zero mean, then this type of shock (uncertainty) on average should cancel out and not have any major effect on the discount rate. In these cases, assumptions about the distribution of stochastic terms are key drivers of the main results.

Similarly, Gollier (2008) allows the possibility of catastrophic recessions into the model.
which has shown to reduce the discount rates. He does this by introducing a Markov process of the consumption growth uncertainty. However, additional assumptions on the preference function are required for the discount rate to decline for all time horizon. Weitzman (2012) assumes a Muth-Kalman Hidden-State Stochastic Process on the consumption growth rate where its uncertainty grows over time horizon. This type of process is also shown to reduce the social discount rates.

Numerous papers have estimated various reduced form econometric models that describe certainty-equivalent discount rates (e.g Gollier et al., 2008; Hepburn et al., 2009; Newell and Pizer, 2007; Groom et al., 2007, 2015). Notably, Newell and Pizer (2003) estimate reduced-form time series process, using two centuries of data on bond yields for the U.S as a proxy for the social discount rates. By assuming interest rates to follow an Auto-Regressive process, they allow the social discount rates to be uncertain and dependent on the persistence parameter of the shocks. When the persistence parameter is equal to 1 (i.e Random Walk), the discount rate declines from 4 percent after the first year to 1 percent after 100 years. Meanwhile, if the persistence parameter is 0.96, the discount rate decreases more gradually, from 4 percent to 3.6 percent over a span of 100 years.

Groom et al. (2007) use the same data set as Newell and Pizer’s (2003) but employ more flexible reduced form models. The first two models are the Random Walk and Mean Reverting models, the same as those analyzed by Newell and Pizer (2003). In addition, they consider the Auto-regressive Integrated Generalized Auto-regressive Conditional Heteroskedasticity (AR-IGARCH) model which allows the conditional variance of the interest rate to vary over time. The fourth model analyzed is the Regime Switching model where interest rates are allowed to shift randomly between two states that differ in mean and variance. The fifth model is the State Space model, permitting the degree
of mean reversion and the variance of the process to change over time. Consistent with previous studies, the social discount rates are shown to be highly sensitive to assumptions about the persistence of the shocks. For instance, the discount rates under the Space State model decline at a much faster rate than those under the Random Walk model.

As noted earlier, numerous authors have employed Regime-Switching Models to describe the discount rates (e.g. Bansal and Zhou, 2002; Gray, 1996; Gollier et al., 2008; Groom et al., 2007; Hamilton, 1988). In these studies, each regime is assumed to have a different mean, variance, and level of persistence in the process. For instance, Gollier et al. (2008) implemented a simulation methodology under the Regime-switching framework, using data from 9 different countries (Africa, Australia, France, Germany, India, Japan, Canada, U.K, and the U.S). They found significant differences in the discount rates across various countries. In particular, France was observed to have produced the sharpest declining discount rates whereas South Africa’s rates showed the slowest rate of decline over time horizon.

In practice $\rho$ is not an observable variable. Several papers have estimated $\rho$ based on the behavior of individuals in lab experiments. These studies have shown that people have a tendency to put higher weight (lower utility discount rate) in the distant future (e.g. Benzion et al., 1989; Loewenstein and Elster, 1992). Some claim that $\rho$ should be purely based on ethical considerations. For instance, Stern (2006) has argued that it should be set equal to the probability of human race extinction which he claims to be 0.01. According to Frank Ramsey, it “$[\rho > 0]$ is ethically indefensible and arises merely from the weakness of the imagination$[.]$” (1928, p543), implying a zero value of $\rho$ while others like Nordhaus (2007) have supported a positive value based on
a descriptive approach. As a result, forming a consensus about $\rho$ (highly subjective and unobservable) has been proven to be difficult. This issue has contributed to one of the most critical problems with discounting future benefits and costs and may be the reason why “no consensus now exists, or for that matter has ever existed, about what actual rate of interest to use” (Weitzman 2001, p260).

Looking at it slightly differently, Weitzman (2001) formulates the question as the problem of aggregating the social discount rates when people have ideological differences in the time preferences. He conducts a survey by asking over 2000 economists with Ph.D. level education around the world about their preferred discount rate. He shows the distribution of the reported rates closely follows a Gamma distribution and that the aggregated social discount rate decreases in people’s ideological differences.

Jouini et al. (2008) utilize Weitzman’s data set to calibrate their model and reach a similar conclusion that different rates of time preference and beliefs about the future state of the economy lead to a lower social discount rate when the elasticity of marginal utility of consumption is larger than 1. In a theoretical framework, Li and Lofgren (2000) formulate the problem from the perspective of a social planner who wishes to maximize a weighted sum of people’s utilities when there are two types of individuals in the society, one with utilitarian and the other with conservationist utility streams whom also differ in ideological preferences for $\rho$. In this setup, the socially optimal weighting scheme is to give dominant weights to individuals with a lower discount rate. In a related paper, Gollier and Zeckhauser (2005) show that if individuals have a preference for different values of $\rho$ that is constant over time, then the social discount rates that maximize the collective welfare will be non-constant and vary over time. In an empirical study, Warner and Pleeter (2001) have found that individual $\rho$ can vary up

\footnote{In section 2.3.3, we provide some implications of the slope of the social discount rate with respect to the pure rate of time preference when the latter is equal to zero.}
Hu and McKitrick (2013) consider a model in which a government planner may seek to distort optimal consumption growth path by distorting the true discount rate. Their framework embeds natural resource stock and political environment which impact the benefits and costs associated with discount rate distortion. For example, a higher initial capital stock is shown to increase the likelihood of discount rate distortion (upwards), thus leading to higher short-term increase in consumption, but it leads to lower long-run growth. On another hand, a decrease in autonomy in policy making is shown to have the opposite effect. These results are consistent with why some resource rich countries suffer from a “resource curse” while others like Canada do not suffer from it. Finally, their results are robust to the introduction of technological change.

Overall, the literature provides a rich set of theoretical and empirical properties of the social discount rate. In particular, it finds that $\rho$ is a critical component of the social discount rate (though there is no clear consensus on how it should be measured), and it is typically assumed unit slope of $r$ with respect to $\rho$. In this paper, we provide in-depth analysis of relaxing this assumption in Ramsey equation. In what follows, we first consider a general model framework, followed by a closed form solution case. Some empirical results are also provided.

2.3 The sign of $\frac{\partial r}{\partial \rho}$ in the general model

2.3.1 Derivation of Ramsey equation

We begin by formally introducing Ramsey growth framework for deriving the social
discount rates. There is a social planner who has perfect knowledge about the future of infinite time horizon. Acting as a trustee on behalf of both present and future generations, his objective is to maximize the intertemporal discounted social welfare function

\[
\max_{\{c(t)\}} \int_0^\infty e^{-\rho t} U(C(t)) dt
\]

subject to

\[
\dot{K}(t) = F(K(t)) - C(t)
\]

where \(\rho > 0\), equation (24) describes the motion of capital (investment), \(K(t)\) and \(C(t)\) are capital stock and consumption level at time \(t\), \(F(\cdot)\) and \(U(\cdot)\) represent production and utility functions respectively. Both \(F(\cdot)\) and \(U(\cdot)\) are assumed to be strictly increasing, twice differentiable and strictly concave. Also assume the usual Inada conditions: \(\lim_{C \to \infty} \frac{\partial U}{\partial C} = 0, \lim_{C \to 0} \frac{\partial U}{\partial C} = \infty\) and \(\lim_{K \to \infty} \frac{\partial F}{\partial K} = 0, \lim_{K \to 0} \frac{\partial F}{\partial K} = \infty\).

Corresponding current value Hamiltonian is given by,

\[ H = U(C(t)) + \lambda(t)(F(K(t)) - C(t)) \]

with \(K(0) = K_0\) given and \(\lambda(t) = \lambda e^{\rho t}\).

From the first-order conditions of the optimal control theory, one can easily derive the following consumption discount rate, also known as Ramsey discount formula

\[
r(t) = \rho + \eta(t) \frac{\dot{C}(t)}{C(t)}
\]

where \(r(t)\) is the social discount rate at time \(t\) (also equal to the marginal product of capital ) and \(\rho\) is the pure rate of time preference (also known as utility discount
rate). Finally, $\eta(t)$ and $\frac{C'(t)}{C(t)}$ are the elasticity of marginal utility of consumption and the consumption growth rate at time $t$. Note that $r(t)$ is the social discount rate at different points in time.

### 2.3.2 Endogenous response to changes in $\rho$ and general equilibrium value of $r$

Equation (25) implies that increases (decreases) in $\rho$, $\frac{C'(t)}{C(t)}$ and $\eta(t)$ increase (decrease) marginal product of capital (the social discount rate). While the growth rate of consumption and the elasticity of marginal utility of consumption influence the marginal product of capital, they do not influence $\rho$ since it is an exogenous parameter in the model. In contrast, $\rho$ is a parameter governing the relative weights of utilities across time so that it will presumably affect the optimal values of endogenous variables, namely the consumption growth rate and the elasticity of marginal utility of consumption. For instance, an increase in $\rho$ implies that if a social planner is faced initially with a situation where consumption is constant over time and if he could increase consumption in any year, then the optimal action would be to choose consumption increase early rather than late. It has not been shown in such a case the effects of the endogenous response of consumption (therefore the consumption growth rate) to changes in $\rho$ and how these would impact the social discount rates. Herein, we define $r$ as the general equilibrium social discount rate where the optimal consumption growth rate is a function of $\rho$. As a result, the slope of $r(t)$ with respect to $\rho$ will depend on how the consumption growth rate and the elasticity of marginal utility of consumption behave to $\rho$ and these effects will be fully captured in $r(t)$ (in our general equilibrium framework).
In order to fully characterize this effect, we first differentiate equation (25) with respect to $\rho$,

$$\frac{\partial r(t)}{\partial \rho} = 1 + \frac{\partial \eta(t)}{\partial \rho} \frac{\dot{C}(t)}{C(t)} + \eta(t) \frac{\partial \dot{C}(t)}{\partial \rho}$$

which can be further expanded to

$$\frac{\partial r(t)}{\partial \rho} = 1 + \left[ \frac{\partial \eta(C(t))}{\partial \rho} \dot{C}(t) + \eta(C(t)) \left( \frac{\partial C(t)}{\partial \rho} C(t) - \frac{\dot{C}(t) \partial C(t)}{C(t)^2} \right) \right]$$

(26)

where the terms in square brackets are the indirect effects of $\rho$ (the endogenous response). Firstly, direct effect arises from the fact that equation (25) is a direct function of $\rho$. Secondly, the indirect effects of $\rho$ are from changes in the optimal consumption when $\rho$ changes which, in turn, may alter $\frac{C(t)}{C(t)}$ and $\eta(t)$ in equation (25).

With some algebraic manipulation and assuming constant elasticity of marginal utility of consumption for simplicity, we get the following simplified expression of equation (26),

$$\frac{\partial r(t)}{\partial \rho} = 1 + \left( \frac{\ddot{\eta}}{\rho g(t)} \right) \left[ \varepsilon_{g(t)} - \varepsilon_{C(t)} \right]$$

(27)

where $g(t)$ is the growth rate of consumption at time $t$, $\ddot{\eta}$ is constant elasticity of marginal utility of consumption, $\varepsilon_{C(t)}$ and $\varepsilon_{g(t)}$ are the elasticities of consumption and the growth rate of consumption with respect to $\rho$ respectively at time $t$.

It is interesting to observe that an increase in $\rho$ does not necessarily imply an increase in the social discount rate. As equation (27) shows, the signs and relative sizes of $\varepsilon_{C(t)}$ and $\varepsilon_{\frac{\dot{C}(t)}{C(t)}}$ are key determinants for both the sign and the size of $\frac{\partial r(t)}{\partial \rho}$. For example, if $\varepsilon_{C(t)} > 0$, $\varepsilon_{\frac{\dot{C}(t)}{C(t)}} > 0$ and $\varepsilon_{\frac{\dot{C}(t)}{C(t)}} > \varepsilon_{C(t)}$ then $\frac{\partial r(t)}{\partial \rho} > 1$ but if $\varepsilon_{\frac{\dot{C}(t)}{C(t)}} < \varepsilon_{C(t)}$ then $\frac{\partial r(t)}{\partial \rho} < 1$ and may even be less than zero depending on the relative sizes of those elasticities (there are total of 8 different cases to consider).
In sum, in the case of constant elasticity of marginal utility of consumption the sign of \( \frac{\partial r(t)}{\partial \rho} \) depends on signs and sizes of \( \varepsilon C(t) \) and \( \varepsilon C(t) \). If we impose absolutely no restrictions on the parameters, then \( \frac{\partial r(t)}{\partial \rho} \) is given by equation (26) which shows that its sign depends on the behaviour of consumption and the consumption growth rates to changes in \( \rho \) and their interactions. So unless all terms in the square brackets amount to zero, a unit slope between the social discount rate and \( \rho \) will no longer hold true which is a common assumption in the literature (e.g. See Dasgupta, 2008; Goulder and Williams III, 2012; Guo et al., 2006). In other words, our results suggest that the impact of \( \rho \) on the social discount rate may be more complex than previously thought in the literature.

Since we were not able to obtain specific results in the general framework such as how the model’s parameters would impact the marginal effect of \( \rho \) on \( r(t) \) (impossible without the functional form of \( \frac{\partial r}{\partial \rho} \)), we resort to a closed-form solution case of the underlying model and analyze its equation (27).

### 2.3.3 A closed form solution model

In general equation (25) is a second order, non-linear differential equation. Now let us assume that production and utility functions have the following form,

\[
\begin{align*}
\text{Production function is } & \quad Y = K(t)^{\alpha} \\
\text{Utility function is } & \quad U = \frac{C(t)^{1+\eta}}{1-\eta}
\end{align*}
\]

with \( \alpha \in [0, 1] \), \( \eta \in [0, \infty] \) and \( t \) denotes time index. Applying the functions above in equation (25), we get the following,

\[
\alpha K(t)^{\alpha-1} = \rho + \dot{C}(t) \frac{\dot{C}(t)}{C(t)}.
\]

\( \text{(28)} \)
and the motion of capital is
\[ \dot{K}(t) = F(K(t)) - C(t). \]  
(29)

Two boundary conditions are: initial capital stock \( K_0 \) and transversality condition 
\[ \lim_{t \to \infty} e^{-\rho t} C(t) - \frac{1}{\alpha} K(t) = 0. \] Equation (29) assumes zero depreciation of capital.

The standard steady-state values of consumption and capital stock are found by setting 
\( \dot{C}(t) = 0 \) and \( \dot{K}(t) = 0 \) in equations (28) and (29). Solving for steady state values we get,
\[ K_{ss} = \left( \frac{\rho}{\alpha} \right)^{\frac{1}{1-\alpha}} \]
\[ C_{ss} = \frac{\rho}{\alpha} K_{ss}. \]  
(30)

As in Smith (2006), applying necessary assumption, \( \eta = \frac{1}{\alpha} \), for tractability, the solutions to equations (28) and (29), subject to \( K_0 \) given and transversality condition are given by
\[ K(t) = \left[ K_{ss}^{1-\alpha} + (K_0^{1-\alpha} - K_{ss}^{1-\alpha})e^{-(1-\alpha)\frac{\rho}{\alpha}t} \right]^{\frac{1}{1-\alpha}} \]  
(31)
\[ C(t) = \frac{\rho}{\alpha} K(t). \]  
(32)

Hence the social discount rate along the optimal growth path is given by
\[ r(t) = \alpha K(t)^{\alpha - 1} \]  
(33)

where \( K(t) \) is given by equation (31).

Now differentiating equation (33) with respect to \( \rho \) yields (using equation (31))
\[ \frac{\partial r(t)}{\partial \rho} = \frac{\alpha^2(1 - e^{-(1-\alpha)\frac{\rho}{\alpha}\rho t})}{\rho^2(\frac{\alpha}{\rho} + (K_0^{1-\alpha} - \frac{\alpha}{\rho})e^{-(1-\alpha)\frac{\rho}{\alpha}\rho t})^2} > 0 \quad \forall \ t > 0. \]  
(34)
Equation (34) is positive for all \( t \) except for \( t = 0 \) where it is also equal to zero and \( t \) denotes time index. The relationship between \( r(t) \) and \( \rho \) is not only shown to be dependent on the level of \( \rho \) itself but also depends on other parameters of the model such as \( K_0 \) and \( \alpha \) as well as time index \( t \). This form is a direct result of the optimal consumption growth path being a function of these parameters. Since \( r(t) \) is a function of the optimal consumption growth according to Ramsey equation, \( \frac{\partial r(t)}{\partial \rho} \) will depend on those parameters of the model as well. For example, increase in \( K_0 \) decreases the value of equation (34) since it only appears in the denominator. Other variables like \( \alpha \) enter more complex in equation (34), appearing non-linearly in both numerator and denominator. This result is apparent from our model set-up: \( \rho \) affects the optimal consumption growth path which, in turn, affects the savings and investment decisions. This inherent trade-off between consumption and investment is captured by equation (29), also known as the motion of capital. By construction \( r(t) \) is defined as the marginal productivity of capital at time \( t \), so that \( r(t) \) will be a function of \( \rho \) with its slope given by equation (34).

The choice of \( \rho \) in Ramsey equation is a contentious one. This issue arises because it is an unobservable and highly subjective variable. Some like Stern (2006) argue based on the normative view that, \( \rho \) should be equal to zero on ethical grounds. Others like Nordhaus claim this rate should be based on market estimation. If we take the limit of equation (34) as \( \rho \) approaches 0\( ^+ \), both numerator and denominator become zero (zero divided by zero). Therefore, by applying the L’Hospital’s Rule, we find its limit as \( \rho \) approaches zero to be infinity. The limit value indicates that even a smallest increase in \( \rho \) increases \( r(t) \) by infinity. In other words, equation (34) is undefined if we were to assign \( \rho = 0 \) based on ethical grounds.

In the following Figures 2.1, 2.2 and 2.3, we show the optimal time paths of capital stock, the marginal product of capital, and the slopes (equation (34)) for three values of \( \rho \). For simplicity we assume \( K_0=1 \).
Figure 2.1. The optimal capital stock over time for three values of $\rho$.

Figure 2.2. The optimal values of $r(t)$ over time for three values of $\rho$. 
Figure 2.3. Equation (34) (Marginal effect of $\rho$ on $r(t)$) for three values of $\rho$ over time.

Figure 2.1 shows the optimal paths of capital stock. For three values of $\rho$ considered, the capital stock increases then reaches a steady state. Notice that higher $\rho$ is associated with the capital stock path that is lower. This result is intuitive from our set up. Higher $\rho$ implies that a representative agent has become more impatient so that his optimal consumption path will have more consumption in the near period compared to previous $\rho$ (given the assumed preference function). Furthermore, according to the motion of capital given by equation (29), this implies that less consumption can be invested in capital in each period resulting in a lower capital stock in the future, ultimately leading to a lower capital stock in the long run. Next, Figure 2.2 depicts the optimal values of $r(t)$ over time. Given the usual form of production function, the $r(t)$ (marginal product of capital) is computed by raising the capital stock at the time to the power of constant $(1-\alpha)$, and this whole term is multiplied by a constant $(\alpha)$. This explains why $K(t)$ increases while $r(t)$ decreases over time but both reach their steady states at the same time.

Figure 2.3 shows $\frac{\partial r(t)}{\partial \rho}$ (we refer to it as the slope) for three values of $\rho$ considered.
Unlike the time paths of the capital stock, the slopes increase until it reaches the peak then descends gradually towards one. Also higher $\rho$ shifts down the slopes for all values of $t$ which is indicative of reduced marginal effect of $\rho$ on $r(t)$ (i.e. lower value of the slope) for higher capital stock. Eventually, both the capital stock and the slope reach their steady states. The long-run capital stock depends on $\rho$ as explained above and the slope converges to a value of one regardless of parameter values as it reaches the steady state. Another interesting result of Figure 2.3 is the different rates of convergence of the slopes for three values of $\rho$. It takes much longer time for the slope to converge to its steady state, lower is the value of $\rho$. In other words, not only the level of $\rho$ impacts the marginal effect of $\rho$ on $r(t)$, it also impacts the speed of convergence towards the long run.

In Figures 2.1 and 2.3, $K(t)$ reaches its steady state very quickly compared to the corresponding slope. The gap in the rate of convergence may seem puzzling at first glance since the slope depends on $K(t)$. For example, after $t = 5$ the capital stock is barely changing, but the slope undergoes a big drop. The difference in the convergence rates, however, become apparent when we look at the meaning and functional form of each variable. The slope measures the changes in the marginal product of capital to changes in $\rho$ at time $t$ and $K(t)$ measures the capital stock at time $t$. If time index affects the slope and the capital stock differently, then there is no reason for both variables to converge into their steady states at the same rate. So this means even when the capital stock is in its steady-state (no change in the capital stock), the slope could still be on its way to the steady state (changing in time). This is precisely why $K(t)$ and $r(t)$ reach their steady state at the same time, but $\frac{\partial r(t)}{\partial \rho}$ reaches its steady state at a different point in time. This is evident when we examine equations (31) and (34) where the latter represents the more complex function of the model’s parameters as well as the time index appearing in both numerator and denominator. As a result, the model’s parameters and time index may have entirely different effects on $K(t)$ and $r(t)$.

In sum, we have shown in this section that the marginal effect $\rho$ on $r(t)$ is no longer equal to one when consumption growth is not assumed constant. This relationship is
directly linked to capital accumulation since there is an inherent trade-off between consumption and current investment in capital in each period which ultimately determines the future capital stock. However in the long run, equation (34) converges to a value equal to one. Two things are worth mentioning. First, this convergence is due to the fact that long-run equilibrium is defined where changes in capital and consumption are equal to zero. Furthermore, this implies that equation (34) by construction will always be equal to one in the long run. In other words, it is the underlying assumptions in the model that derive the results. A second issue is how long a unit of time means in human years. Simply put, a unit of time, in theory, could mean a millennium in practice, so that even if we believed long run assumptions to be reasonable, the notion of a unit of time may have a completely different meaning in reality. It may well be that our world is only half way to the steady state, which means equation (34) is not equal to one; however, in this case, we know it is greater than one. Note that our results in this section depend heavily on the function form of preference function where we knew, even before solving the model, the optimal consumption had to satisfy the 'consumption smoothing' property. For this reason, the results here are conservative in a sense that the consumption growth path is least likely to be sensitive to changes in $\rho$ compared to if we were to adopt other types of preference functions.

Next we present some empirical evidence on $\frac{\partial r(t)}{\partial \rho}$.

### 2.4 Empirical estimates of $\rho$, $\eta$ and $\frac{\partial r}{\partial \rho}$ for the United States

#### 2.4.1 TIME-VARYING SEMI-PARAMETRIC SMOOTH CO-EFFICIENT MODEL
Equation (25) can be used as an empirical model to estimate $\rho$. But $\eta(t)$ is not constant so we need to allow for time-varying parameters. We employ the Time-Varying Semi-Parametric Smooth Coefficient model (denoted TVSSC). Specifically designed to deal with time-series data, this model belongs to a more general class of Semi-Parametric Smooth Coefficient models.

Before getting into the specifics of our methodology, it is worth mentioning a few highlights of the model. First unlike the usual parametric estimators, the unique feature of TVSSC is that it makes no assumptions about the functional form of the coefficients. As a result, we do not have to worry about the usual bias in estimated parameters in TVSSC caused by model misspecification that is otherwise worrisome in other models. This feature is especially useful in our context since in this section we attempt to estimate two parameters that are unobservable in practice, namely $\rho$ and $\eta$, which would presumably mean that their functional forms are also unknown.

The simplest (the usual) way to estimate $\rho$ and $\eta$ in Ramsey formula (equation (25)) is to use Ordinary Least Squares estimator of the following,

$$r(t) = \hat{\beta}_1 + \hat{\beta}_2 g(t) + \epsilon(t), \quad t = 1, 2, \ldots, T$$

where $g(t)$ is the interest rate at time $t$, $\hat{\beta}_1$ and $\hat{\beta}_2$ are estimate of $\rho$ and $\eta$ respectively. Finally, $g(t)$ is the consumption growth rate at time $t$. There is a critical assumption embedded in this framework. It is assumed that both estimated parameters are constant over time. This assumption in our context is highly inappropriate, as noted earlier, since both parameters are not only highly subjective but also unobservable in practice. As a result, if the functional form is misspecified, estimation results on a parametric model can yield misleading conclusions.
The TVSSC can overcome inherent limitations of OLS estimators. No functional form of parameters is assumed and in the context of estimating equation (25), parameters are allowed to vary smoothly over time. This means that we can also estimate the response of \( r(t) \) with respect to \( \hat{\rho}(t) \), \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \). This method is considered semi-parametric since the model contains a finite dimensional parameter of interest but also contains some unknown functions. The approach taken here has been unprecedented in the literature on the social discount rates and to our knowledge, this is the first paper to apply this estimator. We believe some of our discoveries shed light on the issues surrounding highly debated parameters, namely \( \hat{\rho}(t) \), \( \hat{\eta}(t) \) and the value of \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \).

The TVSSC is as follows,

\[
r(t) = X'(t)\delta(Z) + \varepsilon(t), \quad t = 1, 2, ..., T
\]

(35)

where \( r(t) \) denotes the marginal product of capital, \( T \) is the sample size, \( X(t) \) denotes a \( p \times 2 \) vector of explanatory variables of interest, \( \delta(Z) \) is a vector of unspecified smooth functions of \( Z \) and \( t \) (on \( X \)) denotes transpose. In our context \( Z = \frac{t}{T} \). Essentially time is normalized to 1 so that the value of \( Z \) increases monotonically with time. Robinson (1989) showed that this normalization is necessary to ensure the asymptotic justification of the nonparametric smoothing function.

As in Ozturk and Stengos (2014), the vector of smooth functions \( \delta(\cdot) \) can be estimated using a local least squares of the following form,

\[
\hat{\delta}(Z) = [(T* h^q)^{-1} \sum_{j=1}^{T} X(j)X'(j)K(\frac{Z(j) - Z}{h})]^{-1} \ast [(T* h^q)^{-1} \sum_{j=1}^{T} X(j)r(t)K(\frac{Z(j) - Z}{h})]
\]

where \( K(\cdot) \) is a kernel function, \( h = h_T \) is the smoothing parameter for sample size \( T \),
\(j\) is the time index, \(Z(j)\) denotes a \(q \times 1\) vector of other exogenous variables and \(q\) is a constant.

Relating equation (35) to Ramsey equation, \(r(t)\) is the interest rate (proxy for marginal product of capital), \(X\) is a matrix of \(T \times 2\) where the first column is a vector of 1’s and the second column contains the consumption growth rates. The estimated matrix \(\hat{\delta}(Z)\) is a \(T \times 2\) matrix with the first column containing the time-dependent \(\hat{\rho}(Z)\) and the second column containing the time-dependent \(\hat{\eta}(Z)\). Finally, our sample size is \(T = 85\).

### 2.4.2 Data

We use U.S annual time-series data from years 1930 to 2014 to estimate equation (35). For the dependent variable, \(r(t)\), we use market real interest rates for long-term government and corporate bonds. Newell and Pizer (2003) compiled interest rate data from 1796 to 1999 based on Homer and Syllas’s History of Interest Rates (1998). I obtained their data and updated it to 2014. However, the final sample size was restricted by the availability of data on independent variable consumption growth rate, which runs from 1930-2014. This data was obtained from the official website of the Federal Reserve Bank of St. Louis. Newell and Pizer used long term government bonds. I could not find same series so instead, used long term corporate bonds from the website [http://www.measuringworth.com](http://www.measuringworth.com) (herein MW). Finally, we merged those two data sets (Ours and Newell and Pizer (2003)) for the dependent variable, \(r(t)\), for a total of 85 observations (1930-2014).

A few words on merging two data sources: Updating the original Newell and Pizer’s
data from 2000 to 2014 has the benefit of increasing the sample size. However, when doing so, interest rates from two different sources must be compatible. First is the sources of nominal interest rates. For instance, it would be absurd to combine two nominal interest rate sets, one measured in 90-day commercial paper rate and the other in overnight lending rate. Second, the real interest rate, calculated by subtracting the inflation rate from the nominal rate, can be quite different depending on how it is calculated. This is because the inflation rate is a relative measure of prices that depend on the chosen base year.

The nominal interest rates of Newell and Pizer (2003) and MW are long-term government bonds and long-term corporate bonds respectively. Though these are not the same, both rates should be closely related. Next, the CPI (consumer price index) used to compute the inflation rates in Newell and Pizer sets the year 1967 as the base year, whereas, MW uses the average CPI of 1982-1984. It is possible to transform the CPI (of different base years) such that both series have the same base year. However, since the real interest rates between two sources from 1990 to 1999 were relatively close, we felt it was adequate to simply merge the two sources for 2000 to 2014: When we took the difference in the real interest rates from 1990 to 1999, the absolute difference was at most 1.3 percentage points between the two. In the future however, I plan to find the updated interest rates on long term government bonds (2000 and on) to exactly match the Newell and Pizer (2003) series. Thus for now, we believe that two merged series are reasonably compatible, and allows us to extend the Newell and Pizer definition forward in time.
Figure 2.4. U.S real market interest rates from 1930 to 2014.

Figure 2.5. U.S. annual consumption growth rate from 1930 to 2014.

<table>
<thead>
<tr>
<th>Variable</th>
<th>T</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>consumption growth</td>
<td>85</td>
<td>3.2</td>
<td>3.0</td>
<td>-9</td>
<td>12.4</td>
</tr>
<tr>
<td>real interest rate</td>
<td>85</td>
<td>3.4</td>
<td>0.9</td>
<td>1.5</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Table 2.1. Descriptive summary statistics of consumption growth and real interest rates for U.S. from 1930 to 2014.
2.4.3 Results and Discussion

2.4.3.1 Estimation results

![Graph showing real interest rates, estimated $\hat{\rho}(t)$, and 90 percent confidence intervals of $\hat{\rho}(t)$ from 1930 to 2014.]

Figure 2.6. Real interest rates, estimated $\hat{\rho}(t)$, and 90 percent confidence intervals of $\hat{\rho}(t)$ from 1930 to 2014.

Firstly, TVSSC model proved to have enormous explanation power. Our computed r-squared is 0.949, implying 95 percent of variation is explained by the model. Next, Figures 2.6 presents estimated $\hat{\rho}(t)$. Notice that $\hat{\rho}(t)$ is now expressed as a time-dependent parameter (compared to our notation $\rho$ used in Section 2.3). Of course this is because, as shown in Figure 2.6, it is non-constant over time. In our sample period, $\hat{\rho}(t)$ hovers slightly above 3 percent until around 1960 then ascends sharply, reaching the peak around 1980, then descends to slightly below 3.5 percent. The overall shape of the curve reflects a quadratic functional form of $\hat{\rho}$ in time. Notice a striking similarity between $\hat{\rho}(t)$ and $r(t)$ in Figure 2.6, which is indicative of a positive correlation between the two. While they are correlated, it is not perfect. In other words, we do not observe unit slope between $r(t)$ and $\rho$. Also, the value of $\hat{\rho}(t)$ gets as high as 4.8 percent (1980) and does not dip below 3 percent for the whole sampling period. Indeed, this result provides
strong evidence against the idea that \( \rho \) should be equal to zero on ethical grounds (e.g. Stern (2006)). For instance, Stern's (2006) extremely high estimate of the social cost of carbon is based on \( \hat{\rho}(t) \) being equal to 0.1 percent which is assumed to be constant over time.

Figure 2.6 also shows 90 percent confidence intervals around \( \hat{\rho}(t) \). The confidence interval band contains \( r(t) \) for most of the sample period, further reinforcing the close relationship between the two rates. This result suggests that \( \hat{\rho}(t) \) may be affected by the business cycle. The real interest rate, \( r(t) \), which is equal to the marginal product of capital, is directly linked to economic activity. For instance, during bust (recession) and boom, the economy undergoes great fluctuations in total investments and savings. This, in turn, affects the total future capital stock and the future marginal productivity of capital. In Figure 2.6, \( r(t) \) fluctuates substantially over time as a result of natural business cycles (e.g. the Great Depression in the 1930’s, the oil shock in the 1980s, and most recently, the mortgage market failure in 2009). Since our estimated \( \hat{\rho}(t) \) is closely related to \( r(t) \), this suggests that not only is \( \hat{\rho}(t) \) non-constant over time, but more importantly, it is also affected by the underlying economic conditions.

Next, using these estimated parameter values, we propose a simple way to estimate \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \).

### 2.4.3.2 Estimation of \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \)

Using the estimated time-dependent \( \hat{\rho}(t) \) and \( \hat{\eta}(t) \), we estimate \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} = 1 + \frac{\partial \hat{\eta}(t)}{\partial \hat{\rho}(t)} \frac{C(t)}{\hat{\rho}(t)} + \hat{\eta}(t) \frac{\partial C(t)}{\partial \hat{\rho}(t)} \) by the following two-step procedure. But first, we make use of the following assumption,

\[ [A1] \quad \text{The effect of } \hat{\rho}(t) \text{ on } \frac{\hat{C}(t)}{C(t)} \text{ is constant and linear over time.} \]
The effect of $\dot{\rho}(t)$ on $\dot{\eta}(t)$ is constant and linear over time.

The two-step procedure is as follows:

1) First we estimate $\frac{\partial \dot{\eta}(t)}{\partial \dot{\rho}(t)}$ by running OLS regression of the following form: $\dot{\eta}(t) = \hat{\beta}_1 \dot{\rho}(t) + \text{error}$.

2) Second we estimate $\frac{\partial \dot{C}(t)}{\partial \dot{\rho}(t)}$ by running OLS regression of the following form: $\dot{C}(t)(t) = \hat{\beta}_2 \dot{\rho}(t) + \text{error}$.

Therefore by [A1], $\frac{\partial \dot{C}(t)}{\partial \dot{\rho}(t)} = \hat{\beta}_2$ and by [A2], $\frac{\partial \dot{\eta}(t)}{\partial \dot{\rho}(t)} = \hat{\beta}_1$. The following Figure 2.7 illustrates the estimated $\frac{\partial r(t)}{\partial \dot{\rho}(t)}$ over time.

![Figure 2.7](image)

Figure 2.7. Estimate of $\frac{\partial r(t)}{\partial \dot{\rho}(t)}$ for U.S. from 1930 to 2014.

<table>
<thead>
<tr>
<th>Variable</th>
<th>T</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>dr/drho</td>
<td>85</td>
<td>0.84</td>
<td>0.097</td>
<td>0.63</td>
<td>1.16</td>
</tr>
</tbody>
</table>

Table 2.2. Descriptive summary statistics of $\frac{\partial r(t)}{\partial \dot{\rho}(t)}$ for U.S. from 1930 to 2014.

As shown in Figure 2.7, there is some variability in $\frac{\partial r(t)}{\partial \dot{\rho}(t)}$ in time. Initially, the slope reaches a value as high as 1.16 to as low as 0.63 with the mean of 0.84; thus, on aver-
age, one unit increase in $\dot{\rho}(t)$ increases the social discount rate by 0.84. This finding is interesting alone because it means that the slope of $r(t)$ with respect to $\dot{\rho}(t)$ is not equal to one (unit slope) as commonly assumed in the literature. In this simple analysis, it suggests that endogenous response, that is, the effect $\dot{\rho}(t)$ has on $\dot{\eta}(t)$ and $\frac{C(t)}{C(t)}$ to be approximately equal to -0.14. In the general equilibrium framework discussed in Section 2.3.2, this is the case where $\varepsilon \frac{C(t)}{C(t)} < \varepsilon C(t)$ and both terms are positive.

Though this may seem small, it may have a significant impact on the Cost-Benefit Analysis (as discussed in Section 2.2). Let us illustrate this point using the following example. Suppose a regulator faces a dilemma about the choice of $\rho$. For the purpose of our illustration, let us consider the simplest case of all. A regulator is deciding between two values of $\rho$: 3% or 4% (assume these values are constants for simplicity). Also, assume that the corresponding discount rate for the higher value of $\rho$ (4 percent) is equal to 5 percent. This means if he decides to pick $\rho$ equal to 3 percent instead, a decrease in $\rho$ by a percent, its corresponding discount rate would be 3.13 percent, not 3 percent under the usual assumption. To give an idea of what this may imply about the social cost of carbon, Guo et al. (2006) have shown that a 1 percent decrease in the discount rate (from 1% to 0%) caused the social cost of carbon to increase from $13 per ton (in 2005 prices) to $71 per ton (in 2005 prices).

Looking ahead, we would like to estimate $\frac{\partial r(t)}{\partial \rho(t)}$ for countries with different income levels because the initial capital stock was shown to be an important parameter in the closed form solution case. This exercise would require acquiring panel data sets. It would be interesting to estimate the parameters of our interest by different income groups and controlling for various socioeconomic factors. From what we saw from a closed-form solution model, we conjecture that estimated parameters might behave quite differently than for the U.S. We leave this for a future research topic.
2.5 Conclusions

Despite the complicated relationship between the pure rate of time preference, \( \rho \), and the social discount rate, \( r(t) \), the literature often assumes a unit slope between the two, in which a one percent increase in \( \rho \) increases \( r(t) \) by the same amount. This treatment is equivalent to ignoring all the indirect effects of \( \rho \) in Ramsey equation and treating the consumption growth rate and the elasticity of marginal utility of consumption to be constants.

Firstly, in the general theoretical framework where we impose no functional forms on the production and preference, we have shown that the value of \( \frac{\partial r(t)}{\partial \rho} \) can be equal to 1, greater than 1, less than 1 and may even be negative, depending on the relative sizes and the signs of elasticities of marginal utility of consumption and the growth rate of consumption with respect to \( \rho \). It is equal to one (the standard case) only under restrictive conditions, when those elasticities are identical in absolute sizes and have the same signs (i.e. either both positive or both negative). We also found several cases in which an increase in \( \rho \) decreases the social discount rate. Second, in a closed form solution model we showed that \( \frac{\partial r(t)}{\partial \rho} > 0 \ \forall \ t \), but its magnitude depends on many parameters of the underlying model such as \( \alpha \) and initial capital stock \( (K_0) \) as well as its level of \( \rho \). This dependence arises from the inherent trade-off that exists between consumption and investment in the model.

Finally, we employed U.S data to empirically estimate \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \) using a two-step procedure. Note that in the denominator, the pure rate of time preference is expressed as a time-dependent parameter. This form is based on our estimated \( \hat{\rho}(t) \), which is non-constant over time. First \( \hat{\rho}(t) \) and \( \hat{\eta}(t) \) are estimated via the Time-Varying Semi-Parametric Smooth Coefficient Model. We then use these parameter values to estimate \( \frac{\partial r(t)}{\partial \hat{\rho}(t)} \). The results of the first step were somewhat unusual. For our sample period (1930-2014) we found strong non-linearity for \( \hat{\rho}(t) \) where it closely followed a quadratic form.
in time. The estimated value of $\frac{\partial r(t)}{\partial \hat{\rho}(t)}$ is equal to 0.84 for the United States, suggesting the endogenous response in Ramsey equation arising from one unit change in $\hat{\rho}(t)$ to be -0.16.

In sum, this paper has demonstrated the relationship between the pure rate of time preference and the social discount rate to be much more complex than previously thought in the literature, and assuming a unit slope between $r$ and $\rho$ is highly inappropriate in some cases.
Chapter 3

Discounting when The Consumption Growth Path and The Rate of Time Preference are Uncertain and Latter Affects the Former: A Simulation Approach

3.1 Introduction

In Cost-Benefit Analysis, the choice of social discount rate is a contentious one. This issue was accentuated by the emergence of long-lived policies and projects in the last several decades such as carbon taxes and construction of nuclear waste repository sites like the one in Yucca Mountain in Nevada, U.S.A. For example, many critics (e.g Nordhaus (2007b)) of Stern Review (2006) argue that his extremely high estimation of the social cost of carbon stems from the unusually low social discount rate used in the analysis.

The purpose of this paper is to examine the certainty-equivalent social discount rates, herein denoted \( r^c \), when both the pure rate of time preference and the future consumption growth path are uncertain. Our main results are strikingly different from standard results in the literature, that introducing uncertainty causes \( r^c \) to increase instead of lowering it. First, we impose uncertainty in the pure rate of time preference that is constant in discounting horizon. Second, the pure rate of time preference is assumed to change in discounting horizon. Finally, a closed form solution case is examined where
the pure rate of time preference affects the optimal consumption growth path. Guo et al. (2006), herein G6, have derived schedules of \( r^c \) when the future consumption growth rates are drawn from certain distributions (uniform and normal). However, three (implicit) assumptions are made throughout the paper. First, the pure rate of time preference is assumed to be a constant. Second, the pure rate of time preference is considered to have no uncertainty. Third, the consumption growth rate is treated to be independent of this rate. We extend the framework of G6 and examine the behavior of \( r^c \) when one or more of these assumptions are relaxed. This extension allows us to replicate their findings and also be able to put comparisons with the results of our own.

The second chapter of this thesis (previous chapter) specifically examines the relationship between the pure rate time preference and the social discount rate in a general equilibrium framework. The Ramsey equation is given by

\[
r(t) = \rho + \eta \frac{\dot{C}}{C}
\]

where \( \rho \) is the pure rate of time preference, \( \eta \) is the elasticity of marginal utility of consumption and \( \frac{\dot{C}}{C} \) is the optimal consumption growth rate. The theoretical framework of chapter 2 of this thesis suggests that derivation of the social discount rate with respect to \( \rho \) is equal to one only under certain conditions. This is due to the fact that changes in \( \rho \) may also influence the optimal consumption growth rate (i.e. \( \frac{\partial \frac{\dot{C}(\rho)}{\dot{C}}}{\partial \rho} \neq 0 \)). As will be discussed in the literature review section, the majority of papers in the discounting literature assume zero correlation between \( \rho \) and the optimal consumption growth. In this sense, all previous research prior to chapter two of this thesis can be regarded as partial equilibrium approach. We call this case the restricted approach in this chapter.

In the case of the closed form solution model, we have shown that in the previous chapter, this relationship is highly sensitive to parameters of the underlying model such as the initial capital stock, implying that initial income level is one of important factors
in the relationship between the two rates. Using the Time-Varying Semi-Parametric Model, the time paths of $\rho$ and the elasticity of marginal utility of consumption are estimated for the United States. Our approach is related to theirs in a sense that we examine (certainty-equivalent) social discount rates in a general equilibrium framework (unrestricted framework) as well as in a partial equilibrium framework (restricted framework). We also adopt some of their empirical results regarding the functional form of $\rho$ which is discussed in Section 3.4.1.

The combination of uncertainties in optimal consumption growth and in $\rho$ would be fruitful in unpacking a few key relationships that have been left unexplored. Our analysis shows that in general, uncertainty in $\rho$ shifts up the schedule of $r^c$ but its magnitude is crucially dependent on the persistence of the shocks in $\rho$. This result is particularly important because numerous papers have pointed out that uncertainty in the consumption growth (e.g. Weitzman (2012)) or in the discount rate itself (e.g. Newell and Pizer (2003); Groom et al. (2007)) reduce $r^c$. Instead, our results have shown that its effects depend on the source of the uncertainty and its assumed properties.

Thus, we have shown that identification and decomposition of different types of uncertainty in Ramsey equation are crucial to proper valuation of discounted future benefits and costs, especially for projects and policies with long-life spans.

The paper is organized as follows. Section 3.2 discusses the past literature on uncertainty and the social discount rates. Section 3.3 discusses the type of consumption uncertainty analyzed in this paper and derivation of $r^c$. Section 3.4 discusses the properties of $\rho$ and its uncertainty. Section 3.5 outlines the Monte-Carlo simulations for the restricted case where the consumption growth is independent of $\rho$ and reports the results. Section 3.6 presents a closed form solution case and the results. Finally, Section 3.7 is the conclusion.

3.2 Literature review
The concept of social discount rate was first discussed more than a century ago, in a seminal 1889 work by Bohm-Bawerk. However, it was several decades later, in 1928, when the first formal theoretical derivation of the social discount rate was published by Frank Ramsey. His famous work is known as the Ramsey equation which relates the social discount rate to the pure rate of time preference (showing the innate nature of humans to consume more now than later) and the underlying consumption growth. Numerous papers have emerged since then that extend the original Ramsey framework. For example, many papers have examined the effects of consumption uncertainty on the social discount rate schedule (e.g. Mankiw, 1981; Gollier, 2002a, 2002b, 2008; Weitzman, 2012). Pindyck and Wang (2012) have modeled the Ramsey framework that incorporates unlikely but catastrophic events such as wars and natural epidemics. More recently, several papers have put forward arguments for using 'dual discounting’ when utility is a function of various consumption goods, rather than of general composite goods (e.g. Baumgartner et al., 2015; Hoel and Sterner 2007; Weikard and Zhu, 2005).

Though these extensions reveal invaluable characteristics about the social discount rate, a major result in the literature which originated from Ramsey (1928) is that, regardless of the underlying scenarios, the value of discount rate chosen by a social planner (hence social discount rate) is some function of the representative agent’s pure rate of time preference.

There are two major strands of literature on the theme of uncertainty and discounting. The first strand assumes social discount rate as a whole to be uncertain (see Newell and Pizer, 2003, Groom et al., 2007). The second strand, which is our approach to modeling uncertainty, assumes uncertainty in the consumption growth rate in Ramsey equation (see Mankiw, 1981; Gollier 2002a, 2002b, 2008; Weitzman, 2012).

Newell and Pizer (2003) assume that though there is considerable uncertainty in the social discount rates, past values are informative about the future rates. They estimate various reduced form econometric models that describe $r^c$ as being comprised of two parts: A permanent and a random component. Specifically, they examine how $r^c$
behaves in relation to different structures of the random component that are persistent over time. Using long-term U.S real interest rate data, they run simulations that include random walk and mean-reverting structures on the random component of $r^c$. The results show $r^c$ is highly sensitive to both the structure and the persistence of the shocks. For instance, $r^c$ under the random walk structure on the random component declines at a much faster rate over time than the mean-reverting structure.\(^9\)

Groom et al. (2005) replicate the results of Newell and Pizer, even using the same data sets, but add three additional models. First, they add Auto-regressive Integrated Generalized Auto-regressive Conditional Heteroskedasticity (AR-IGARCH) model which allows the conditional variance of the interest rate to vary over time. Second, they add Regime Switching Model (RSM) where interest rates are allowed to shift randomly between two states that differ in mean and variance. Third, State Space model is employed, allowing the degree of mean reversion and the variance of the process to change over time. In a related paper, Gollier et al. (2008) empirically estimate time-dependent schedules of $r^c$ using real interest rates (as a proxy for discount rates) of Government bonds for France, Japan, India, South Africa, Australia, Canada, Germany, U.K and the U.S. The estimated $r^c$ for all sample countries exhibit a similar pattern in that they all decline smoothly over time. However, there are major differences as well. In particular, slopes of $r^c$ with respect to time across countries are quite different. For instance, interest rates of France showed the sharpest decline while the South African rates showed the slowest decline over time. However, they offer no real explanation for cross-sectional differences in $r^c$ across countries.

In a theoretical framework, Weitzman (2012) models the consumption growth rate as an unobservable random walk. A unique feature of this model is that consumption uncertainty grows with the time horizon. In other words, a margin of error in prediction for consumption growth rate in the year 2060 is assumed to be higher than for the year 2040. The results show that $r^c$ declines quadratically over discounting horizon.

\(^9\)Note that the discount rate here is defined as the marginal rate between two adjacent periods (i.e. between $t$ and $t+1$), so that the time argument is over time, as oppose to being over discount horizon.
G6, whose framework we follow closely in this paper, examine the case in which uncertainty is discrete and embedded in the consumption growth rates. Assuming that realized consumption growth rates for each discounting horizon are drawn from a certain distribution (uniform and normal), first certainty-equivalent discount factors are computed, and the corresponding $r^c$ are then calculated. However, their analysis assumes $\rho$ is fixed without uncertainty and does not affect the optimal consumption growth path. We extend their framework to incorporate and analyze uncertainty in $\rho$ in restricted and unrestricted frameworks.

Next, we fully discuss the framework of G6 and our approach to modeling uncertainty in $\rho$ and its properties.

### 3.3 Uncertainties in the consumption growth rates and the pure rate of time preference: The restricted approach

In this section, we outline derivation of $r^c$ for discrete consumption uncertainty. Adopting the framework of G6, let us consider an economy in which all future consumption growth rates are uncertain. For every discounting horizon, a consumption growth rate is realized from a set containing three values with equal probability. We refer to this case as the restricted approach. In Section 3.6, we consider a case in which consumption growth depends on the value of $\rho$ which we refer to as the unrestricted approach. Let $g_1$, $g_2$ and $g_3$ represent three growth rates and $p_1 = p_2 = p_3 = \frac{1}{3}$ are the respective probabilities. We adapt the discount factor form proposed by Loewenstein and Prelec (1992) so the certainty-equivalent discount factor (herein $CF(H)$) is
calculated as the following:

$$CF(H) = \frac{1}{3} \left( \frac{1}{(1 + r_1(H))^H} + \frac{1}{3} \frac{1}{(1 + r_2(H))^H} + \frac{1}{3} \frac{1}{(1 + r_3(H))^H} \right)$$ \hspace{1cm} \text{(36)}$$

where $H$ denotes time horizon and

$$r_i(H) = \rho(H) + \eta g_i(H) \quad i = 1, 2, 3.$$ \hspace{1cm} \text{(37)}$$

Equation (37) is the standard Ramsey equation where $\eta$ is the elasticity of marginal utility of consumption, $\rho(H)$ is the pure rate of time preference (we consider two types of $\rho(H)$ and this is fully discussed in the subsequent section) and $g_i(H)$ is one of three possible consumption growth rates. Note that for simplicity, we have suppressed all other parameters in equations (36) and (37) except $H$.

From equation (36) the corresponding $r^c(H)$ is given by the following formula,

$$r^c(H) = \left( \left( \frac{1}{CF(H)} \right)^\frac{1}{H} - 1 \right) \times 100\% \quad \forall H$$ \hspace{1cm} \text{(38)}$$

where equation (38) represents the certainty-equivalent discount rate for time horizon $H$, not the instantaneous period-to-period rate at time $H$ in the future. As noted in Guo et al. (2006), values of 'average' and 'instantaneous' rates may differ but should on the same principle behave similarly. Thus our approach can be interpreted as a comparison of average $r^c(H)$ across various discounting horizons $H$ (or policies with different life spans).

### 3.4 Properties of $\rho(H)$

We assume $\rho(H)$ in equation (37) is comprised of two parts: A permanent (with
zero variance) component and a random component. This is described by the following equation,

\[ \rho(H) = \hat{\rho}(H) + \varepsilon(H) \forall H \]  

(39)

where \( \hat{\rho}(H) \) and \( \varepsilon(H) \) are permanent and random components respectively. We consider two types of permanent component in our Monte-Carlo simulations and this is outlined in the following section 3.4.1. The random component, \( \varepsilon(H) \), is an autocorrelated shock to the pure rate of time preference. Equation (39) means the following. There is a consensus (among the citizens per se) about the value of \( \hat{\rho}(H) \) but there is uncertainty because this value may change in discounting horizons. This is due to random shocks so that the overall realized value of \( \rho(H) \) for discounting horizon \( H \) is computed by adding up the permanent part and the random part. The following sections, 3.4.1 and 3.4.2, provide detailed discussion on each component.

### 3.4.1 Properties of permanent component \( \hat{\rho}(H) \) in \( \rho(H) \)

Two types of permanent components are as follows:

- **Type 1**: \( \hat{\rho}(H) = \rho \forall H \)
- **Type 2**: \( \hat{\rho}(H) = \rho_0 + \rho_1 H + \rho_2 H^2 + \rho_3 H^3 \) where \( \rho_0, \rho_1, \rho_2 > 0 \) and \( \rho_3 < 0 \).

Type 1 assumes \( \hat{\rho}(H) = \bar{\rho} \forall H \) where its rate is constant for all discounting horizon. For Type 2, \( \hat{\rho}(H) \) follows a polynomial form in \( H \). This case was motivated by empirical results from chapter 2 of this thesis, where we employed Time-Varying Smooth Coefficient model and U.S data to estimate \( \rho \). In fact, the parameters in Type 2, \( \rho_0, \rho_1, \rho_2 \) and \( \rho_3 \) are chosen to match the properties of the estimated \( \rho \) in chapter 2. There are two advantages of employing Type 2. First, we can examine how \( r^c(H) \) behaves in comparison to non-constant \( \rho(H) \). Second, even if the actual property of \( \hat{\rho}(H) \) is given by Type 1, we can examine how \( r^c(H) \) responds to changes in \( \hat{\rho}(H) \) that is constant in
$H$. In this sense, this case can also be thought of as the analysis of the slope of $r^c(H)$ with respect to $\rho$ (i.e. $\frac{\partial r^c(H)}{\partial \rho}$).

3.4.2 Uncertainty (random component) in $\rho(H)$

We assume the random term $\varepsilon(H)$ is persistent and defined by the following equation,

$$\varepsilon(H) = \delta\varepsilon(H - 1) + (\varphi(H))^2. \quad (40)$$

where $\delta \in [0, 1]$ is a parameter governing the persistence of shocks between discounting horizon $H$ and $H - 1$. Note that this does not imply $H$ is uncertain, but it describes how error terms are generated for different $H$. For instance, a value of $\delta$ close to one in equation (40) means that the value of error is likely to be high for $H$ if the error term for $H - 1$ was high and vice versa. Note that this does not imply that errors are cumulating for higher discounting horizons, instead, it implies that errors are correlated between nearby discounting horizons. \(^{10}\) Finally, $\varphi(H)$ denotes the shock for $H$ which is drawn from the following distribution,

$$\varphi(H) \sim \mathcal{N}(0, 0.04), \quad (41)$$

where it is normally distributed with a mean of zero and variance of 0.04. Note that though we have chosen a particular mean and variance, the overall results of the Monte-

\(^{10}\) In the literature, the variable $H$ may carry a different meaning depending on how the certainty-equivalent discount rate is defined. If the certainty-equivalent rate is defined as the instantaneous period-to-period rate in to the future, then $H$ in equation (40) needs to be defined as the time index. On the other hand, if the certainty-equivalent rate is defined to be the 'average' rate for time horizon $H$ as it is in our approach, then $H$ in this case, denotes discounting horizon.
Carlo simulations are robust to changes in these parameter values. As noted earlier, the role of $\delta$ in equation (40) is to control for persistence in shocks, which turns out to be one of the key parameters governing the shape of $r^c(H)$. We show the resulting $r^c(H)$ for various values of $\delta$ in section 3.5.2. Finally, in order to avoid negative values we square randomly drawn shock in equation (41) in equation (40). Hence, the squared term in equation (40) is distributed with Chi Square.

3.4.3 General properties of $r^c(H)$ with respect to uncertainties in the consumption growth rate and $\rho$

The variance of consumption growth can be written as

$$\theta^2_g = \frac{\sum_{i=1}^{N}(g_i - \bar{g})^2}{N}$$

where $N$ is the number of possible consumption growth rates (in equation (36), $N$ is equal to three), $g_i$ is the $i^{th}$ growth rate and $\bar{g}$ is the average growth rate. Solving above for $g_i$ we obtain the following,

$$g_i = \sqrt{N \cdot \theta^2_g - \sum_{-i}^{N}(g_i - \bar{g})^2 + \bar{g}}.$$

Plugging in the equation above in equation (3) we get\textsuperscript{11},

$$r^c(H) = \left[ \sum_{i=1}^{N} \frac{1}{p_N \left( \rho + \eta \left( \sqrt{N \cdot \theta^2_g - \sum_{-i}^{N}(g_i - \bar{g})^2 + \bar{g}} \right) \right)^\pi} \right]^{\frac{1}{\pi}} - 1$$

where $p_N$ is the probability of $i^{th}$ consumption growth rate being drawn. For simplicity,\textsuperscript{11}Note that in equation (38), $N = 3$. The $r^c(H)$ expressed here is for any values of $N$.\textsuperscript{67}
we assumed a uniform distribution so that there is equal probability on all of the possible (discrete) growth rates. All other parameters are defined as before. From above, it is not clear whether an increase in $\theta^2_g$ increases or decreases $r^c(H)$. For instance, an increase in $\theta^2_g$ may be accompanied with an increase $N$ (possible growth rates) but it could also increase or decrease $\bar{g}$. In other words, in order to pin down the slope of $r^c(H)$ with respect to $\theta^2_g$, we need to know how it affects $\bar{g}$. Figure 1a in the following illustrates a case in which increase in its variance lowers $r^c(H)$.

![Figure 3.1](image.png)

Figure 3.1. The effects of increase in consumption growth uncertainty in $r^c(H)$. In this example, increase in its uncertainty lowers $r^c(H)$.

Analysis of the variance of $\rho$ (denoted $\theta^2_\rho$) in $r^c(H)$ is a much more difficult task. In fact, it is impossible to derive a functional form of the slope of $r^c(H)$ with respect to $\theta^2_\rho$. However, we attempt to provide some useful insights into this issue. If we assume that $\theta^2_\epsilon(H) = \theta^2_\epsilon(H-1)$, then by equation (40) and (41) it follows that $\theta^2_\epsilon = \frac{\text{var}(\varphi(H))}{1-\delta^2}$. Furthermore, by equation (39), $\theta^2_\rho \approx \theta^2_\epsilon = \frac{\text{var}(\varphi(H))}{1-\delta^2}$. Note that as $\delta$ approaches one, $\theta^2_\epsilon$ reaches infinity. In other words, the more persistent the error in between discounting horizons, the higher the variance of $\rho$ will be. Unfortunately, this is as far we can discuss about the properties of $\rho$ as there is no analytical way to express $r^c(H)$ as a function of $\theta^2_\rho$. 

68
(i.e. $r^c(\sigma_r^2)$). We leave this for the next section, where we examine $r^c(H)$ for various underlying cases in the Monte-Carlo simulation framework.

3.5 Monte-Carlo simulations for the restricted framework

We formulate Monte-Carlo simulations for four various cases that differ in properties about the pure rate of time preference (constant vs. non-constant in discounting horizon) and its persistence in shocks. Except in the benchmark case, we introduce stochastic component in $\rho(H)$ according to equations (40) and (41). For each case, we simulate 1000 schedules of discount rates over various discounting horizons then compute the average discount rate schedule. We summarize each case in detail in section 3.5.2.

3.5.1 Step-by-step Monte-Carlo simulation procedure

First, generate random sample $\{\varepsilon(H)\}_{H=1}^{100}$ for random error component in equation (40) and compute $\{\hat{\rho}(H)\}_{H=1}^{100}$ given parameter values. Second, compute the Ramsey discount rate according to equations (37) and compute its $CF(H)$ using equation (36). Next, compute $r^c(H)$ according to equation (38). We repeat steps (1) - (2) for 1000 times and compute the average values of $r^c(H)$ for all $H$. We repeat steps (1)-(3) for two different values of persistent parameter $\delta$ in equation (40).

3.5.2 Results
Table 3.1: Descriptions of cases 1-7 in Figure 3.2 (cases 1-3) and Figure 3.3 (cases 4-7).

<table>
<thead>
<tr>
<th></th>
<th>Color code</th>
<th>Permanent component of rho</th>
<th>TC**</th>
<th>PE ***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case 1</td>
<td>black</td>
<td>fixed and equal to 0.03</td>
<td>no</td>
<td>n/a</td>
</tr>
<tr>
<td>case 2</td>
<td>red</td>
<td>fixed and equal to 0.03</td>
<td>yes</td>
<td>0.3</td>
</tr>
<tr>
<td>case 3</td>
<td>green</td>
<td>fixed and equal to 0.03</td>
<td>yes</td>
<td>0.8</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case 4</td>
<td>*</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>case 5</td>
<td>black</td>
<td>function of H (case 4)</td>
<td>no</td>
<td>n/a</td>
</tr>
<tr>
<td>case 6</td>
<td>red</td>
<td>function of H (case 4)</td>
<td>yes</td>
<td>0.3</td>
</tr>
<tr>
<td>case 7</td>
<td>green</td>
<td>function of H (case 4)</td>
<td>yes</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Note: *case 4 shows the $H$ dependent pure rate of time preference described in section 3.4.1 as

Type 2. ** Transitory component of rho ***Persistence in errors (PE) is denoted by a greek letter delta in equation (40).

Figure 3.2. $r^c(H)$ when $\hat{\rho}$ is constant in discounting horizon (Type 1).
Figure 3.3. $r^c(H)$ when $\hat{\rho}$ is non-constant in discounting horizon (Type 2).

Figure 3.2 shows $r^c(H)$ when $\rho(H)$ is constant in discounting horizon (Type 1). The black curve is our benchmark case where $\rho(H)$ has no uncertainty. This case is the same as G6 where the average $r^c(H)$ is declining smoothly over $H$ in the presence of the consumption uncertainty. Next, we introduce shocks to $\rho(H)$ that is persistent in discounting horizon. For the red curve, the persistence parameter is assigned a value of 0.3 which means that 30 percent of the shock is carried over between discounting horizons. Notice that this shifts up the whole schedule of $r^c(H)$, indicating uncertainty in $\rho(H)$ increases $r^c(H)$ for all $H$.

For the green curve in Figure 3.2, a value of 0.8 is assigned to the persistence parameter ($\delta$) which means 80 percent of shock is carried over between discounting horizon. Notice that the green curve sits above the red curve for all values of $H$. This result indicates that the effect of uncertainty in $\rho(H)$ is exacerbated by the increase in persistence of the shocks. In other words, the higher the persistence of shocks in $\rho(H)$, the greater the effect of its uncertainty in $r^c(H)$. This outcome is analogous to results of Newell
and Pizer (2003) which show that for uncertainty to have a profound effect, its shocks must be persistent over time (however they assume the social discount rate as a whole to be uncertain).

In fact, our results are obvious and intuitive when we examine the way $r^c(H)$ is calculated. Given the assumptions about consumption uncertainty, discount factor in equation (36) is decreasing in $r(H)$. Also, notice that $\rho(H)$ appears in all $r_i(H)$ where $i = 1, 2, 3$. In turn, this implies that positive shocks to $\rho(H)$ that are persistent in $H$ will decrease the discount factor. As a result, $r^c(H)$ given by equation (38) shifts up in the presence of persistent uncertainty in $\rho(H)$.

Figure 3.3 depicts a case in which $\rho(H)$ follows a polynomial shape in $H$ (Type 2). The general properties of the effects of uncertainty in $\rho(H)$ are the same as above. Its uncertainty raises $r^c(H)$ for all $H$, and higher persistence exacerbates this effect. One result is worth mentioning. Comparing the pink curve with the black curve, notice the gap between the two curves is not constant in $H$. This implies that $\frac{\partial r^c(H)}{\partial \rho(H)}$ is not equal to one and is $H$ dependent.

It is possible to generate many schedules of $r^c(H)$ by varying the parameters in the model. The values chosen here are to provide a graphical illustration showing that under a range of possible parameter values and structures for $\rho(H)$, $r^c(H)$ can vary widely. In fact, it is possible to generate cases where $r^c(H)$ declines at a much faster (or slower) rate than ones depicted in Figures 3.2 and 3.3.

Next, we simulate a closed form solution of equation (38). A major advantage of this framework over restricted analysis is that we no longer need to assume an independent relationship between $\rho$ and the optimal consumption growth path, as the latter is
determined endogenously by the model.

3.6 The pure rate of time preference dependent consumption growth path and \( r^c(H) \): The unrestricted approach

3.6.1 A closed form model

To examine the effects of uncertainties on \( r^c(H) \) when the consumption growth path is affected by the pure rate of time preference, we resort to a closed form solution to Ramsey equation. Let us assume production and utility functions to have the following form,

Production function is \( Y = K(H)^\alpha \)

Utility function is \( U = \frac{C(H)^{1-\eta}}{1-\eta} \).

As in Smith (2006) assuming \( \eta = \frac{1}{\alpha} \), \( \dot{K}(H) = F(K(H)) - C(H) \) (motion of capital) and \( \lim_{H \to \infty} e^{-\rho H}C(H)^{-\frac{1}{2}}K(H) = 0 \) (transversality condition), we obtain the following consumption growth function,

\[
g(\rho, H, K_0, \alpha) = \frac{\dot{C}(\rho, H, K_0, \alpha)}{C(\rho, H, K_0, \alpha)} = \alpha(\alpha[K_{ss}^{1-\alpha} + (K_{0}^{1-\alpha} - K_{ss}^{1-\alpha})e^{-((1-\alpha)\frac{H}{\alpha})}]^{-1} - \rho) \quad (42)
\]

where \( K_{ss} = (\frac{\rho}{\alpha})^{\frac{1}{\alpha-1}} \), \( K_0 \) is the initial capital stock, \( \rho \) is the pure rate of time preference and \( \alpha \) is the technology parameter. In contrast to our previous case where we assumed the consumption growth to be exogenously given (i.e. \( \rho \) did not influence the consumption growth path), the solution to the closed form model given by equation (42) shows it not only depends on \( \rho \) but is also a function of other parameters of the model such as
Thus, uncertainty in \( \rho \) will affect \( r^c(H) \) in two ways. First is the direct effect since Ramsey equation is a direct function of \( \rho \). Second is the indirect effect where its uncertainty affects the optimal consumption growth path, thereby influencing the value of \( r^c(H) \) indirectly. The chapter 2 of this thesis thoroughly discusses this indirect effect without uncertainty. In this section, we analyze these effects in Monte-Carlo simulation framework.

Note that the consumption growth uncertainty is modeled slightly different than our previous case (restricted framework). We suppose that the consumption growth rate based on equation (42) (given all parameter values) can fluctuate around its value by three percentage points. The reason for modeling uncertainty in this way is because the consumption growth rate is determined endogenously by the closed form model as opposed to being exogenously given. Uncertainty in \( \rho \) is modeled the same as before where properties of random components are described by equations (39), (40) and (41). Note that we only consider Type 1 where \( \rho(H) \) is constant for all values of \( H \).

Hence in each period we have the following discount rate possibilities with equal probabilities,

\[
\begin{align*}
    r_1(\rho, H, K_0, \alpha) &= \rho + \eta \ast (g(\rho, H, K_0, \alpha) - 0.03) , \quad p_1 = \frac{1}{3} \\
    r_2(\rho, H, K_0, \alpha) &= \rho + \eta \ast (g(\rho, H, K_0, \alpha) , \quad p_2 = \frac{1}{3} \\
    r_3(\rho, H, K_0, \alpha) &= \rho + \eta \ast (g(\rho, H, K_0, \alpha) + 0.03) , \quad p_3 = \frac{1}{3}
\end{align*}
\]  

(43)

where \( \rho \) is given by equation (39) and \( g(\cdot) \) is given by equation (42).

As before, \( CF(\rho, H, K_0, \alpha) \) is calculated by multiplying the discount factor with corresponding probabilities

\[
CF(\rho, H, K_0, \alpha) = \frac{1}{3} \ast \frac{1}{(1 + r_1(\rho, H, K_0, \alpha))^H} + \frac{1}{3} \ast \frac{1}{(1 + r_2(\rho, H, K_0, \alpha))^H} + \frac{1}{3} \ast \frac{1}{(1 + r_3(\rho, H, K_0, \alpha))^H} \forall H
\]  

(44)
where \( r_1(\cdot) \), \( r_2(\cdot) \) and \( r_3(\cdot) \) are given by equation (43). \( r^c(\cdot) \) is then calculated using equation (38). For simplicity, we suppress all parameters in equations (44) and (38) except \( H \) in the following results section.

### 3.6.2 Monte-Carlo simulation steps for a closed form solution model

First, generate the random sample \( \{\varepsilon(H)\}_{H=1}^{100} \) for the random error component in equation (40) and compute \( \{\hat{\rho}(H)\}_{H=1}^{100} \) given parameter values. Second, compute the Ramsey discount rate according to equations (43) and compute its \( CF(H) \) using equations (44). Third, compute \( r^c(H) \) according to equation (38). Next, repeat steps (1) - (2) 1000 times and compute the average values of \( r^c(H) \) for all \( H \). Then repeat steps (1)-(3) for two different values of persistent parameter \( \delta \) in equation (40). Finally, do steps (1)-(4) for two different values of \( \rho \).

### 3.6.3 Results of a closed form model

<table>
<thead>
<tr>
<th>Figure 3.4 case 8</th>
<th>black</th>
<th>fixed and equal to 0.05</th>
<th>no</th>
<th>n/a</th>
</tr>
</thead>
<tbody>
<tr>
<td>case 9</td>
<td>green</td>
<td>fixed and equal to 0.05</td>
<td>yes</td>
<td>0.3</td>
</tr>
<tr>
<td>case 10</td>
<td>red</td>
<td>fixed and equal to 0.05</td>
<td>yes</td>
<td>0.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 3.5 case 11</th>
<th>blue</th>
<th>fixed and equal to 0.03</th>
<th>no</th>
<th>n/a</th>
</tr>
</thead>
<tbody>
<tr>
<td>case 12</td>
<td>orange</td>
<td>fixed and equal to 0.03</td>
<td>yes</td>
<td>0.3</td>
</tr>
<tr>
<td>case 13</td>
<td>purple</td>
<td>fixed and equal to 0.03</td>
<td>yes</td>
<td>0.8</td>
</tr>
</tbody>
</table>

**Transitory component of rho (TC)***Persistence in errors (PE) is denoted by a greek letter delta in equation (40).

75
Figure 3.4 illustrates three cases depicted in different colors. First, the black curve is our benchmark case where $\rho$ is equal to 0.05 (constant in $H$ and no uncertainty).
$r^c(H)$ increases initially then gradually decreases in $H$. Next, we introduce persistent uncertainty in $\rho$. The green curve represents a case in which the persistence parameter is assigned a value of 0.3 ($i.e \delta = 0.3$ in equation (40)). Notice that the green curve sits above the black curve for all values of $H$. This shift-up implies that uncertainty in $\rho$ increase the values of $r^c(H)$ for all discounting horizons. Next, the red curve corresponds to case 10 where the persistence parameter is equal to 0.8. Consistent with the results obtained in the restricted framework, higher persistence is shown to further shift up $r^c(H)$, indicating that it exacerbates the influence of the uncertainty effect on $r^c(H)$.

Next, we perform the same Monte-Carlo simulations as before but now decrease the value of $\rho$ from 0.05 to 0.03. All cases are depicted in Figure 3.5 by different colors. The general shape of each curve is the same as before, where $r^c(H)$ increases initially and decreases smoothly in $H$. However, we observe some significant changes in the effects of uncertainty in $\rho$. Notice that introducing uncertainty as well as increasing its persistence does not shift up $r^c(H)$ as it did for $\rho = 0.05$ and all cases considered in the restricted framework. Three curves converge and overlap one another after $H = 17$. The change in the shapes of $r^c(H)$ in regards to uncertainty arises from indirect effects in a closed form framework. Unlike in the restricted framework where the consumption growth rates are exogenously given, the optimal consumption path is dependent on the value of $\rho$ in a unrestricted (closed form) framework. Figure 3.4 shows that indirect and direct effects of uncertainty work against each other in equal magnitude (approximately), resulting in a convergence of $r^c(H)$ for three cases considered. Hence, the level of $\rho$ plays a key role in the magnitude of the indirect effect, as shown in Figures 3.3 and 3.4. This result is significant and adds to the literature since most past papers on the issue of uncertainty are based on a restricted framework where all indirect effects are ignored altogether which may significantly alter the shape of $r^c(H)$ as shown in this paper.
Note that it is possible to generate many schedules of $r^c(H)$ by varying the parameters in the model. Also, our results here hinge on some of the assumptions imposed on the parameters in order to yield a closed form solution. However, our main point here is that the indirect effect, that is the effect of changes in the optimal consumption growth path from uncertainty in $\rho$, plays a key role in the shape of $r^c(H)$; it should not be overlooked.

### 3.7 Conclusions

This paper has examined two types of uncertainties and their effects on $r^c(H)$ for various underlying scenarios in restricted and unrestricted frameworks. Our main findings are as follows. First, the shape of $r^c(H)$ critically depends on the assumptions about whether or not $\rho$ is uncertain and the degree of persistence of its shocks. In particular, we find that introducing uncertainty in $\rho$ in conjunction with consumption uncertainty raises the $r^c(H)$ as compared to the case without uncertainty in $\rho$. In fact, this result is completely opposite to standard results in literature, where uncertainty in consumption growth rate decreases $r^c(H)$. This result is important since it illustrates not all types of uncertainty lowers $r^c(H)$.

Next, we analyzed the effects of uncertainty in $\rho$ in a restricted framework by employing a closed form solution model where the optimal consumption growth path is a function of $\rho$. First, the uncertainty effect of $\rho$ on $r^c(H)$ depends on the level of $\rho$. When the value of $\rho$ is assigned 0.05, we found that results are essentially the same as in the restricted framework where the uncertainty shifts up $r^c(H)$ for all $H$ and increasing the persistence in its shocks further exacerbates this effect. When we reduce $\rho$ to 0.03, however, the effect of uncertainty in $\rho$ only lasts for a short discounting horizon. This is because different values of $\rho$ impact the consumption growth path differently, exacerbating or diminishing the uncertainty effect of $\rho$ on $r^c(H)$ which is completely overlooked in the restricted framework.
Overall, we have demonstrated that uncertainty of $\rho(H)$ and its relationship with the optimal consumption growth rates play a vital role in the overall shape of $r^c(H)$ which is one of the most crucial, if not the most important variable, in Cost-Benefit Analysis.
References


equilibrium.” Econometrica, 82(1), 41-88.


Pindyck, R.S., 2013. "Climate change policy: what do the models tell us? (No. w19244)." National Bureau of Economic Research.


