Introducing Moral Virtue Ethics into Normative Economics for Models with Endogenous Preferences*

Vipul Bhatt[1] Masao Ogaki[2] Yuichi Yaguchi[3]

James Madison University *Keio University* *Chuo University*

Abstract

This paper proposes a framework to balance considerations of welfarism and virtue ethics in the normative analysis of economic models with endogenous preferences. We introduce the moral evaluation function (MEF), which ranks alternatives based purely on virtue ethics, and define the social objective function (SOF), which combines the Social Welfare Function (SWF) and the MEF. We illustrate the application of our framework using two examples of endogenous preferences. First, is the rational addiction model (Becker and Murphy, 1988) and second is the tough love altruism model (Bhatt and Ogaki, 2012). In both examples we show that maximizing only the SWF may not yield a socially desirable state if the society evaluate some preferences to be better than other preferences in terms of moral virtue. This problem can be resolved by using the SOF to evaluate alternative social states.

---

*This paper was prepared in part while Ogaki was a visiting scholar of Institute for Monetary and Economic Studies at the Bank of Japan, and earlier versions were circulated under the title of "A Reformulation of Normative Economics for Models with Endogenous Preferences." We would like to extend our special thanks to Noah Smith for introducing us to Pollak's (1978) work and to the concept of meta-preferences, as well as for many comments and discussions about many aspects of this paper, and to Toyotaka Sakai for his suggestions about the formulation of the social objective function. We would also like to thank Takanori Ida, Mamoru Kaneko, Fumio Ohtake,Julio Rotemberg, Makoto Saito, and seminar and session participants at the Bank of Japan, the City University of Hong Kong, Kyoto University, Xiamen University, James Madison University, the 2012 Association of Behavioral Economics and Finance Meeting, the 2013 Japanese Economic Association Meeting, the Canon Institute for Global Studies Conference on Macroeconomic Theory and Policy 2014, the 2014 Society for Advancement of Economic Theory Conference, and the 16th Marco Conference for comments. Ogaki's research was partially supported by JHPS Kakenhi Grant Number 25285063. Bhatt's research was partially funded by the JSPS Postdoctoral Fellowship for Foreign Researchers.

[1]Bhatt: Assistant Professor, Department of Economics, James Madison University, 421 Bluestone Drive, ZSH 433, Harrisonburg, VA-22807. Email:bhattvx@jmu.edu.

[2]Ogaki: Professor, Department of Economics, Keio University, 612 Mita-Kenkyu-shitsu 2-15-45 Mita Minato-ku, Tokyo 108-8345, Japan. E-mail: mogaki@econ.keio.ac.jp.

[3]Yaguchi: Graduate Student, Graduate School of Economics, Chuo University, Tokyo, Japan. E-mai: a14.3sr5@g.chuo-u.ac.jp

1. Introduction

Many theoretical and empirical studies have emphasized and identified various channels through which preferences might be endogenously determined in the economy. In the models studied in the literature of intergenerational cultural preference transmission and formation (see Bisin and Verdier (2011) for a survey), children's preferences are affected by parents' decisions. Habit formation models have been used in macroeconomics (see, e.g., Lawrence et al. (2005)), and finance (see, e.g., Constantinides (1990)). Addiction models have been used in microeconomics (e.g., Becker and Murphy (1988)). In the literature of behavioral economics, reference points are often endogenously determined (see, e.g., Kőszegi and Rabin (2006)).

In normative economics we seek to evaluate social states. The two widely accepted normative criteria are the Pareto principle and the Bergson-Samuelson social welfare function (SWF, henceforth). The basis for both of these methods is welfarism.[1] There are two main issues in using standard SWF-based welfare analysis when preferences are endogenous. First, preference ordering conditional on endogenous economic variables cannot be used as a yardstick for the evaluation of social states. To compare two social states, we need an exogenous basis for such an evaluation. Second, given that preferences may be numerous, some preferences may be considered "better" in terms of moral virtue. Pollak (1978) introduces the concept of unconditional preference ordering and suggests the use of such an ordering for normative analysis when preferences change endogenously. Pollak's proposal resolves the

[1]Sen (1979) identifies welfarism as "the principle that the goodness of a state of affairs depends ultimately on the set of individual utilities in that state, and-more demandingly-can be seen as an increasing function of that set."

first issue since by definition, the unconditional preference ordering is exogenous. However, it does not address the second issue. Even though the unconditional preference ordering is exogenous, such a criterion is based on purely welfarist considerations and hence cannot rank alternatives in terms of moral virtues. If a society values virtue, we may not want to rely exclusively on unconditional preference ordering in policy evaluation; we may require an evaluative framework that explicitly accounts for moral virtue considerations.

Hence, we desire an explicit accounting for moral virtue in normative economics. According to the Stanford Encyclopedia of Philosophy "Virtue ethics" is currently one of three major approaches in normative ethics. Virtue ethics emphasizes virtues, or moral character, in contrast to approaches that emphasize duties or rules (deontology) or those that emphasize the consequences of actions (consequentialism)."[2,3,4]

In this paper we propose a policy evaluation procedure that balances welfarism and virtue ethics. For this purpose, we first define a *moral evaluation function* (MEF, hence forth) that expresses evaluations based on virtue ethics. We then define an *social objective function* (SOF, henceforth) that weighs both the MEF and the SWF in evaluating alternative social states. Kaplow and Shavell (2001) prove that any social evaluation function that is not pure welfarist will violate the weak Pareto criterion.[5] As a result our proposed SOF will violate the weak Pareto criterion. However, in models with endogenously determined preferences

---

[2] `http://stanford.library.usyd.edu.au/entries/ethics-virtue`

[3] Welfarism is a form of consequentialism.

[4] Sandel (2009), after considering other major alternatives, promotes Aristotle's moral virtue ethics. According to Aristotle, "moral virtue comes about as a result of habit." In his explanation, these are "the virtues we get by first exercising them, as also happen in the case of the arts as well."

[5] Sen (1970) makes similar arguments about weak Pareto criterion and his analysis can be considered a special case of the framework used in Kaplow and Shavell (2001).

such a violation may be natural. To address this issue Bhatt et al. (2015) propose the *Modified Weak Pareto Criterion* that that adds moral virtue considerations and apply it to a model of endogenous altruism.[6] In this paper, we extend their criterion to a more general setting that allows for moral virtues other than altruism. Accordingly, we propose the *Criterion of Moral Virtue Ethics* that can be used to rank conditional preference orderings in terms of purely moral virtue ethics considerations. Further, we propose the *Modified Criterion of Moral Virtue Ethics* that balances the considerations for welfarism and moral virtue ethics in the evaluation of social states. The application of our mathematical framework requires that SWF satisfies the weak Pareto criterion, MEF satisfies the criterion of moral virtue ethics, and SOF satisfies both the modified weak Pareto criterion and the modified criterion of moral virtue ethics.

For the purpose of illustrating our approach, we use two examples of endogenous preferences model. The first example is based on the rational addiction model of Becker and Murphy (1988) and the second example is based on the tough love altruism model of Bhatt and Ogaki (2012). In both examples we wish to make the following two points. First, in models with endogenous preferences, a Pareto improvement in terms of the unconditional preference ordering may not be socially desirable in terms of the conditional preference ordering. Second, a more balanced approach that weighs both welfarism and moral virtues can resolve the problem faced by the Pareto principle. and policy recommendation from such an

[6]Bhatt et al. (2015) adapt the modification of the Pareto criterion suggested by Temkin (2011) for this purpose. The main purpose of the present paper is to propose a mathematical framework in order to introduce virtue ethics in evaluating alternative social states in economic environments with endogenous preferences. Bhatt et al. (2015) adopt our approach, and explore the relationship between our approach with three major ethics theories (consequentialism, virtue ethics, and deontology).

approach can be very different from those resulting from pure welfarism.

As our first example we focus on the rational addiction model (Becker and Murphy, 1988) which is one of the standard approaches for modeling the consumption of addictive goods such as alcohol, cigarettes, binge eating etc. In this framework, an individual chooses the level of addictive good consumption by maximizing his life time utility. The non-zero level of addictive good consumption will result from such maximization as long as the benefit from consumption exceed any cost of future addiction. An important policy implication of this framework is that the welfare maximizing tax rate is zero as long as there are no external costs associated with the consumption of the addictive good.

Using numerical simulations we first show that moving from a positive tax rate on the addictive good to zero tax rate represents a Pareto improvement in terms of the decision-maker's unconditional preference ordering. However, using conditional preference ordering we find that such a change may not be desirable to the decision-maker. For interpreting this result, imagine a society in which an addictive drug such as heroin has been decriminalized and been taxed. For the sake of the simplicity, imagine that there is no externality and that the representative consumer is rationally deciding how much drug to consume. To lower the tax rate will be a Pareto improvement in this situation in terms of the unconditional preference ordering. However, in our example, the consumer with less addiction in the status quo prefers the positive tax rate in terms of his conditional preference ordering. As long as the society evaluates preferences with less addiction to be better than preferences with more addiction, the society may not wish to rely purely on welfarism.

We illustrate how maximizing a SOF that weighs both welfarism (captured by SWF) and moral virtues (captured by MEF) can resolve the aforementioned problem with the Pareto Principle, and can generate very different policy recommendations. In our framework evaluation of conditional preference ordering in terms of virtue ethics is based on a MEF. In this example we regard having zero stock of addiction as a virtue. Hence, we formulate the MEF such that large deviations from this virtue leads to lower values.

We evaluate different tax rates based on maximizing the SOF using numerical simulations and report several findings of interest. First, the Pareto improvement we obtained earlier by moving from a zero tax rate to a positive tax rate leads to a lower value of the SOF, and hence is not socially desirable. Second, the optimal tax policy based solely on the maximization of social welfare (captured by an SOF with zero weight on MEF) is to have a zero tax on the addictive good. Finally, we show that adding moral virtue considerations in the policy evaluation process yields very different policy recommendations. For instance, maximizing an SOF with a small positive weight on MEF is sufficient to generate positive tax on addictive good consumption in our model. This is an important departure from the existing literature. Incorporating moral virtue considerations may imply optimum positive taxation of the addictive good, even when the individual is rational and there are no externalities associated with the addictive behavior.[7]

As our second example we extend the tough love altruism model from Bhatt and Ogaki

[7] Gruber and Kőszegi (2004) investigate the issue of optimal taxation on addictive goods when individuals exhibit hyperbolic discounting and hence have time inconsistent preferences. They find that in such an environment the optimal policy will be to have a positive tax rate for the addictive good. This is in contrast to our framework which generates positive taxation even when an individual's preferences are time consistent (our example assumes exponential time discounting).

(2012), adding a bequest motive for the parent. This induces a trade-off for the parent between childhood transfers and adulthood bequest. The money saved by lowering childhood transfers can be used to increase parental bequest during the child's adulthood. In this setting the government has a policy tool, the bequest tax rate, that can be used to influence the optimizing behaviors of the parent and the child. For instance, a policy that increases the bequest tax rate will reduce the incentive to leave a bequest, and hence would lead to higher parental transfers to the child. This in turn would lower the child's discount factor in our model.

We use numerical simulations and show that a policy that raises the bequest tax rate from negative to zero represents a Pareto improvement in terms of unconditional preference ordering.[8] However, using conditional preference ordering we find that such a change may not be desirable to the child.

To show how our framework can resolve this problem we focus on the virtue of patience in this example. We view the time discount factor as determining the altruism of the present self toward her future self. If the time discount factor is less than one, then the present self is considered too selfish, while if it exceeds one, then the present self is considered to have excessive altruism. Hence, we define the virtue of patience as when the child's discount factor is one. Such a formulation of the virtue of patience is espoused by many economists and philosophers. [9] We formulate the MEF such that large deviations from this virtue yield lower values.

[8]The negative tax rate means that the government subsidizes bequests. As we illustrate in Section 4, a policy to maximize the SWF leads to a positive bequest tax rate but a policy to maximize the SOF can lead to a negative tax rate in order to promote more bequests and higher patience.

[9]In section 4 we further elaborate the rationale for such a formulation.

Using numerical simulations, we illustrate the main predictions of our model economy when changes in the bequest tax rate are evaluated using the SOF. There are several findings of interest. First, the laissez-faire policy of setting the tax rate to zero does not maximize social welfare (i.e., SWF). Second, the SWF is maximized at a positive tax rate, which in our model economy implies that the child's patience is being influenced by the government policy. Given that the policy is already affecting the child's preferences, we argue that it is irresponsible for the government to completely ignore moral virtue considerations. Third, we show that if we select the appropriate weight on the MEF, evaluation of policy based on SOF maximization can resolve the problems we encountered using the Pareto principle. Specifically, the Pareto improvement we obtained by moving from negative bequest tax rate to a zero rate leads to a lower value of the SOF and hence is not socially desirable. Finally, for a given weight on the MEF, the optimum policy may actually be to set the tax rate to be zero. Hence, we believe that the common practice of using the laissez-faire motivation for ignoring moral virtues in normative economic analysis may not be justified, especially when preferences are endogenously determined.

The rest of the paper is organized as follows. Section 2 provides a brief review of the related literature. Section 3 presents our theoretical framework and defines the MEF and SOF. Sections 4 and 5 first highlight the limitation of Pareto Efficiency in policy evaluation and then illustrates the application of our framework for the rational addiction model and the tough love altruism model, respectively. Section 6 concludes.

2. Related Literature

This study emphasizes the need for adding moral value considerations when evaluating social states in economic models with endogenous preferences. A recent study, related to the tough love altruism example we use in this paper, is by Doepke and Zilibotti (2014). They develop a theoretical framework for transmission of preferences between generations where alternative parenting styles (authoritarian, authoritative, and permissive) may emerge in equilibrium depending on parental preferences and socioeconomic environment.[10] A paternalistic parent in their framework cares about his child's welfare and attempts to affect his child's choice either by influencing preferences directly or by imposing restrictions on choice sets. Although the framework proposed by (Doepke and Zilibotti, 2014) is general, their analysis focuses on patience. They argue that a key area of disagreement between parents and children is about delayed rewards due to children innately having a lower level of patience than what is considered desirable by their parents. They use their model of patience to highlight how socio-economic factors affect parenting styles.

In this paper our proposal to develop an MEF represents an effort to provide a mathematical framework for the evaluation of social states using virtue ethics. As such, the concept of the MEF can be viewed as a reply to a call by Sandel (2013) to bring more value judgment into economics. Instead of relying solely on virtue ethics for this purpose, we seek to combine welfarism and virtue ethics using the SOF. In the same issue of the *Journal of Economic Perspectives*, Bruni and Sugden (2013) argue that classical and neoclassical economics

[10] Bisin and Verdier (2011) provide a comprehensive review of the literature on intergenerational cultural transmission of preferences.

already incorporate many elements of virtue ethics when "market virtues" are considered. The virtue of patience, on which we focus here, can be considered a market virtue. Thus, we argue that economics can benefit from formalizing the notion of market virtues with an approach such as ours.

A review of the relatively sparse literature on this topic identifies three approaches to bring moral considerations into economics (see, e.g., Hausman and McPherson (1993) and Goldfrab and Griffith (1991) for surveys). These are: 1) moral values as norms, 2) moral values as constraints on behavior, and 3) moral values as preferences. Our approach is most closely related to the meta-preference framework (see, e.g., Sen (1974, 1977), Hirschman (1984), and George (1984)). Meta-preferences are preferences one may have about one's own preferences or about the preferences of others. For example, imagine a non-voter who wants to vote in order to be a good citizen, or a smoker who does not want to smoke. In both cases there is a meta-preference about the preference itself. Although such meta-preferences most commonly derive from moral values (e.g. the duty of a good citizen to vote, in the above example), it is possible to have a non-moral basis as well (the desire not to smoke for health reasons). Such a view is pertinent to our research question, since meta-preferences can provide a normative guide to cope with the conflict between the manifest choice and what our moral values dictate. In this sense the meta-preference framework is a natural way to incorporate moral value considerations in economic models. Our proposed MEF applies this framework to rank conditional preference orderings in models with endogenous preferences, for the purpose of introducing virtue ethics into this class of models. Our application of this framework is more related to the "sense of duty" emphasized by Sen (1974, 1977) than to the free choice emphasized by George (1984).

An important consideration in incorporating virtues in policy evaluation is their universality. Dahlsgaard et al. (2005) conduct an extensive survey of philosophical and religious traditions in the East and the West using written texts from Confucianism,Taoism, Buddhism, Hinduism, Athenian Philosophy, Judaism, Christianity, and Islam. They found that the following six virtues appeared in these writings: courage, justice, humanity, temperance, wisdom, and transcendence. For the examples of virtue we use in this paper (patience and self-control) the relevant core virtue identified by Dahlsgaard et al. (2005) is that of temperance. They find that in seven out of eight traditions studies by them temperance is explicitly stated as a moral virtue whereas in Confucianism it is thematically implied (see Table 2 on pg 211, Dahlsgaard et al. (2005)). They argue that such strong convergence across varied traditions is indicative of universality of these core virtues and hence allows a non-arbitrary basis for classifying virtuous behavior across traditions.

Our paper is also related to the recent literature on the economics of happiness. (Frey, 2008, p. 5) lists eudaimonia as one of the three concepts of happiness. Eudaimonia is Aristotle's concept of happiness as a "good life," defined by the acquisition of and use of virtue. Hence our MEF can be viewed as an expression of an aspect of eudaimonia. Benjamin et al. (2014) used surveys with personal and policy scenarios to estimate relative marginal utilities. They estimated high relative marginal utilities not only for happiness and life satisfaction, but also for aspects related to values (morality and meaning), among other things. Thus they show that eudaimonic aspects are important for policy considerations. Our framework introduces moral virtues in evaluating alternative social states. One implication of our proposed framework is that policy interventions may be aimed at enhancing character and foster moral virtues. Cunha and Heckman (2007) identify non-cognitive skills to include values such as perseverance, time preference, and self control. Heckman and Mosso (2014) survey the literature on interventions aimed at enhancing cognitive and non-cognitive skills during childhood. They suggest that early childhood interventions have lasting effects, are more effective than programs aimed at helping disadvantaged adolescents, and an important channel through which they improve adult outcomes is the enhancement of non-cognitive skills.[11] More importantly, the findings from this literature suggest that most promising interventions involve active mentoring. They define mentoring to involve teaching values such as perseverance and cooperation among other character values. Our paper is also related to the literature of behavioral normative economics, in which many models explicitly or implicitly have endogenous preferences. For example, the reference point of prospect theory is often simply assumed to be the level of the initial endowment. Because the initial endowment has been determined endogenously in the economic system (represented more generally by a dynamic model), prospect theory implies a model with endogenous preferences. In a companion paper, Bhatt et al. (2015) provide a review of the literature on behavioral normative economics, and highlight the dominance of welfarism as a basis for policy evaluation in this field of inquiry. They also develop a model of endogenous altruism a ĺa Mulligan

[11]For example Heckman et al. (2013) used dynamic factor approach to evaluate the effect of Perry Preschool Program on later life outcomes such as health, wages, and education. They attribute the effects of this program mainly through the improvement of non-cognitive skills.

(1997), in which a worker can devote resources to become more altruistic toward a disabled stranger. In this model they consider the virtue of altruism toward strangers, which is not a market virtue. Hence, the virtue ethics framework can also incorporate non-market virtues into economic models.

3. Reformulating Policy Evaluation based on Virtue Ethics

In this section we propose a framework that explicitly incorporates virtue ethics considerations in normative economic analysis. Our approach is based on three evaluation functions. The first is the social welfare function (SWF), which captures welfarist considerations. The second is the moral evaluation function (MEF), which is based on virtue ethics. Finally, we have the social objective function (SOF,) which weighs both welfarism and virtue ethics. In this section we formalize these concepts and then illustrate their application in the context of the rational addiction model and the tough love altruism model in sections 4 and 5, respectively.

Consider an economy with N agents. Let x denote a social state and $U_i(x)$ be the utility function of agent i, and $\psi_i(x)$ be a function that expresses properties of the endogenous utility function of agent i. Let $SWF(U_1(x), ..., U_N(x))$ be the social welfare function. The moral evaluation function (MEF) is a function $MEF(\psi_1(x), ..., \psi_i(x))$ that evaluates $(\psi_1(x), ..., \psi_i(x))$ in terms of moral judgments such as deviations of these properties from perfect moral virtue. The social objective function $SOF(MEF(x), SWF(x))$ is a function that evaluates social states by considering both moral virtue and welfarism.

Just as the SWF is required to satisfy the Weak Pareto Criterion for pure welfarism, we need formal criteria that add ethical considerations of moral virtue for MEF and SOF. In order to achieve this, we first need a modification of the Weak Pareto Criterion that allows for ethical factors in comparing social states. This is because any social evaluation that is not that not pure welfarist, such as those based on our proposed SOF, will violate the weak Pareto criterion (Kaplow and Shavell, 2001). To address this issue, a companion paper Bhatt et al. (2015) adapts Temkin's modification of the Pareto criterion (Temkin, 2011, p. 408), and propose the *Modified Weak Pareto Criterion: Given two social states x and y, if everyone strictly prefers x to y, then x should be evaluated to be better than y for society as long as x is not evaluated to be worse than y in terms of other ethically relevant factors.* The conditional statement implied by as long as in the aforementioned modified criterion allows for the possibility that ethical considerations such as moral virtue may outweigh purely welfarist considerations.

Second, we need a criterion that can rank conditional preference orderings in terms of purely moral virtue ethics considerations in order to implement the MEF based evaluation proposed by us in this paper. We adapt Bhatt et al. (2015) criterion of Moral Virtue of Altruism to more general *Criterion of Moral Virtue Ethics: Given two social states x and y, if at least one persons conditional preference orderings strictly better in terms of moral virtue ethics and everyone elses conditional preference ordering is at least as good in terms of moral virtue ethics in x than in y, then x should be evaluated to be better.*

Finally, for SOF, we need to modify the above criterion to allow for the possibility that other ethically relevant factors such as welfarism may outweigh the considerations for moral virtue ethics. Hence we define the *Modified Criterion of Moral Virtue Ethics : Given two social states x and y, if at least one persons conditional preference ordering is strictly better in terms of moral virtue ethics and everyone elses conditional preference ordering is at least as good in terms of moral virtue ethics in x than in y, then x should be evaluated to be better as long as x is not evaluated to be worse than y in terms of other ethically relevant factors.*

In our proposed mathematical framework, SWF needs to satisfy the Weak Pareto Criterion, MEF needs to satisfy the Criterion of Moral Virtue Ethics, and SOF needs to satisfy both the Modified Weak Pareto Criterion and the Modified Criterion of Moral Virtue Ethics.

4. Rational Addiction and Virtue Ethics

Consider an individual who derives utility from the consumption of an addictive good (a_t) and a non-addictive good (c_t).[12] The individual also derives utility from the stock of past consumption of the addictive good denoted by S_t. We also assume, as in (Becker and Murphy, 1988), that the period utility of the individual is non-additive separable in a_t and S_t but is additive separable in the addictive good and the non-addictive good. Hence, the period t instantaneous utility is assumed to take the following form:

$$u_t = v(a_t, S_t) + q(c_t) \qquad t = 0, 1 \tag{1}$$

[12]In formulating our example we borrow from the framework used in Gruber and Kőszegi (2004) who introduced hyperbolic discounting into the rational addiction approach. In this paper we abstract from hyperbolic discounting and illustrate that incorporating virtue ethics leads to significant differences in optimal taxation on the addictive good even with exponential discounting.

In the above formulation, the positive cross-partial derivative, $v_{aS}(a_t, S_t) > 0$, indicates the addictive nature of the good as it's consumption will increase the future marginal utility. We assume that the stock of past consumption of the addictive good evolves as follows:

$$S_{t+1} = (1-d)S_t + a_t \quad t = 0, 1 \tag{2}$$

where d if the rate of depreciation of the stock. Let p_t denote the price of the addictive good and the price of the non-addictive good is normalized to 1. Let y_0 denote the exogenously given income in period 0 and b_0 denote the first period savings. For simplicity we assume that there is no second period income and the individual simply consumes his first period savings that earn a gross interest rate of R. We also assume that the consumption of the addictive good is tax at a time-invariant rate denote by τ and the individual receives a subsidy every period denoted by z_t. The budget constraints faced by the individual in each period are given as follows:

$$\begin{aligned} (3) \qquad & Period\ 0: \ p_0 a_0 + c_0 + b_0 = y_0 - \tau a_0 + z_0 \\ & Period\ 1: \ p_1 a_1 + c_1 = R b_0 - \tau a_1 + z_1 \end{aligned}$$

We can combine the above two constraints and write the intertemporal budget constraint as follows:

$$p_0 a_0 + \frac{p_1 a_1}{R} + c_0 + \frac{c_1}{R} = y_0 - \tau(a_0 + \frac{a_1}{R}) + z_0 + \frac{z_1}{R} \tag{4}$$

In our framework, the optimization problem of the individual can be expressed as follows:

$$\max_{c_0, c_1, a_0, a_1} u_0 + \beta_1 u_1 \tag{5}$$

$$\textit{subject to } (4)$$

4.1. Limitation of Pareto Efficiency

In this section we emphasize a key limitation of the Pareto principle in evaluating social states in the rational addiction framework. Following (Gruber and Kőszegi, 2004) we assume the following specification for the addictive and non-addictive components of the utility function:

$$v(a_t, S_t) = \alpha_a a_t + \alpha_S S_t + \alpha_{aa}\frac{a_t^2}{2} + \alpha_{SS}\frac{S_t^2}{2} + \alpha_{as} a_t s_t \tag{6}$$

$$q(c_t) = \alpha_c c_t + \alpha_{cc}\frac{c_t^2}{2} \tag{7}$$

where α_a, α_c, and α_{aS} are positive implying positive marginal utilities for addictive and non-addictive goods. α_S, α_{cc}, α_{aa}, and α_{SS} are assumed to be negative.

We assume that $S_0 = 0$, so that the consumer starts his life without any addiction. The unconditional utility function, which represents her *unconditional preference ordering* is defined by the following expression:

$$U(x) = v(a_0, 0) + q(c_0) + \beta_1(v(a_1, S_1) + q(c_1)). \tag{8}$$

Given a particular value Q for the state variable of the stock of the addictive good, S_1, the conditional utility function, which represents *conditional preference ordering*, is given by the following expression:

$$U(x|S_1 = Q) = v(a_0, 0) + q(c_0) + \beta_1(v(a_1, Q) + q(c_1)). \tag{9}$$

We numerically solve the decision-maker's optimization as a non-linear root finding problem. Table 1 summarizes the results for a given set of parameter values. Suppose we want to assess whether changing the tax on addictive good from $\tau_0 = 0.05$ to $\tau_1 = 0.0$ is optimal. Let $x(\tau_i)$ be the allocation under the tax rate of τ_i, $U(x(\tau_i))$ be unconditional utility under τ_i. Let $U(x(\tau_1)|S_1(\tau_0))$ be the conditional utility given $S_1(\tau_0)$ (the equilibrium stock of addictive good when the tax rate is τ_0.05).

Table 1 presents the optimal values of the unconditional and conditional utilities. We can observe that in terms of the unconditional utility function, we have a Pareto improvement when the policy changes from τ_0 to τ_1:

$$U(x(\tau_1)) > U(x(\tau_0)) \tag{10}$$

Table 1: Pareto Efficiency and Policy Evaluation: Rational Addiction

Global Parameters		
$\alpha_a = \alpha_c = \alpha_{as} = 0.01$; $\alpha_{aa} = 0.01; \alpha_{SS} = -0.001$; $\alpha_{cc} = -0.001$		
$y_0 = 5$; $d = 0.5$; $R = 1.1$; $p_0 = p_1 = 0.5$; $\beta = 0.95$; $S_0 = 0$		
	$\tau_0 = 0.05$	$\tau_1 = 0.0$
$U(x(\tau_i))$	0.0480	0.0482
$U(x(\tau_1)\|S_1(\tau_0))$	-	**0.0449**

However, in terms of the conditional utility function given the original tax rate, the decision-maker is made worse off by this policy change:

$$U(x(\tau_1)|S_1(\tau_0)) < U(x(\tau_0)|S_1(\tau_0)) \tag{11}$$

The results of Table 1 highlight the limitation of the Pareto principle in evaluating policies when preferences are endogenous. If the society values no addiction as a virtue then $\tau_0 = 0.05$ is the socially more desirable policy when compared to the Pareto improvement (in terms of unconditional utility) represented by $\tau_1 = 0.0$.

4.2. Introducing Moral Virtues in Policy Evaluation

We now illustrate the application of our theoretical framework that balances welfarism and moral virtues within the rational addiction framework. For this purpose, we define the SWF to be the same as the unconditional utility function:

$$SWF = v(a_0, 0) + q(c_0) + \beta(v(a_1, S_1) + q(c_1)). \tag{12}$$

The moral evaluation function (MEF) is given by:

$$MEF = -(S_1)^2 \tag{13}$$

The above formulation of the MEF implies that a larger future stock of addictive good is evaluated to be morally undesirable.

For the purpose of defining the SOF we have to account for the fact that MEF and SWF are in different units and hence not directly comparable. To over come this problem we utilize the Nash SWF framework of Kaneko and Nakamura (1979). For this purpose, we first need to define a worst case scenario for both the SWF and the MEF. These formulations are given by:

$$\overline{SWF} = u_0(\overline{a}_0, \overline{c}_0, \overline{s}_0) + \beta_1 u_1(\overline{a}_1, \overline{c}_1, \overline{s}_1) \tag{14}$$

$$\overline{MEF} = -\ (y_0)^2 \tag{15}$$

In the above definition, the worst possible scenario from moral virtue considerations is when all of the resources in the first period are devoted toward the second period's stock of addictive good. For the worst possible scenario from the welfare considerations, we specify a grid for each of the consumption levels, $\{\overline{a}_t, \overline{c}_t\}$ where $t = 0, 1$ and evaluate the SWF for each combination over the grid.[13] We choose the minimum level over this grid for $\overline{SWF}$. The SOF is given by the following expression:

[13]We fix $S_0 = 0$ which implies $S_1 = a_0$ from (2). For each consumption level we specify a grid over $[0, 5]$ that increments by 0.5.

(16) $$SOF = (MEF - \overline{MEF})^{\alpha}(SWF - \overline{SWF})^{(1-\alpha)}$$

The above social objective function combines the concepts of welfarism and moral virtue. Under welfarism the objective is to maximize the SWF. Hence, if $\alpha = 0$, maximizing the SOF is an expression of welfarism. If the objective is to achieve moral virtue, then MEF is the relevant objective function. Hence, if $\alpha = 1$, maximizing the SOF is an expression of moral virtue ethics. Hence, for $0 < \alpha < 1$, maximizing the SOF is an expression of a balanced approach that combines welfarism and moral virtue considerations.

Using the above framework we state our main result in the following proposition:

Proposition 1 *The optimum tax rate on addictive good consumption is strictly positive if, for* $\tau = 0$, $SWF \neq \overline{SWF}$, $MEF \neq \overline{MEF}$, $\dfrac{da_0^*}{dtau} < 0$, *and the consumer's optimization is obtained with interior solutions.*

Proof: See A.1 in the appendix for a proof. ∎

The most important condition for proving the above proposition is that $\dfrac{da_0^*}{d\tau} < 0$. Even though this condition can be violated in general, it should hold for most reasonable specifications of the economy. The particular functional form for the utility function we use in our simulations is commonly employed in the rational addiction literature and for that parameterization $\dfrac{da_0^*}{d\tau} < 0$.

4.2.1. Simulation Results

Using the parametric specification outlined in the previous section we solve the optimization problem of the individual numerically for different values of the addictive consumption tax

Table 2: Optimal Tax Policy- Welfarism vs Moral Virtue

Global Parameters:

$\alpha_a = \alpha_c = \alpha_{as} = 0.01$; $\alpha_{aa} = 0.01; \alpha_{SS} = -0.001$; $\alpha_{cc} = -0.001$

$y_0 = 5$; $d = 0.5$; $R = 1.1$; $p_0 = p_1 = 0.5$; $\beta = 0.95$; $S_0 = 0$

	$\tau = 0$	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$
a_0	1.3775	0.9786	0.5986	0.2377	0.0010
a_1	1.9763	1.5485	1.1408	0.7537	0.4883
c_0	1.6157	1.8259	2.0271	2.2172	2.3450
c_1	1.9770	2.1790	2.3705	2.5535	2.6758
S_1	1.3775	0.9786	0.5986	0.2377	0.0010
$SOF(\alpha = 0)$	**0.2894**	0.2892	0.2887	0.2880	0.2874
$SOF(\alpha = 0.05)$	0.3602	**0.3607**	0.3606	0.3600	0.3593
$SOF(\alpha = 0.3)$	1.0767	1.0893	1.0961	**1.0983**	1.0974
$SOF(\alpha = 0.8)$	9.6209	9.9317	10.1261	10.2202	**10.2343**
$SOF(\alpha = 1)$	23.1025	24.0424	24.6417	24.9435	**25.0000**

rate, τ. In our solution algorithm we impose the government's budget constraint: $z_t = \tau a_t$.

Table 2 presents the results of this exercise for a given set of parameters. We consider three alternative policy considerations. The first is based on welfarism which involves maximizing only the social welfare function (SWF). In our framework this implies setting $\alpha = 0$ and maximizing $SOF(\alpha = 0)$. The second is based solely on moral virtue ethics and aims to maximize the moral evaluation function (MEF). This obtained by setting $\alpha = 1$ in our model. Finally, the third is based on a balanced principle of weighing both welfarism and moral virtue ethics which can be achieved in our framework by setting $\alpha \in (0, 1)$ and then maximizing the social objective function(SOF).

There are several findings of interest. First, a policy that solely maximizes welfare (as captured by $SOF(\alpha = 0)$) sets that optimal tax rate on addictive good at 0. This is consistent with the main prediction of the rational addiction model with no external costs, the optimal tax rate on the addictive good should be zero. Second, even a small weight on moral virtue considerations may yield very different tax implications. For example, for a weight of $\alpha = 0.05$, the optimal policy that maximizes $SOF(\alpha = 0.05)$ is $\tau = 0.05$. Finally, because the economy in Table 2 is the same as the one in Table 1, it is of interest to investigate whether policy evaluation based on the SOF can resolve the limitation of the Pareto principle highlighted earlier. As we observed from the Table 1 simulations, the change in the tax rate from $\tau_0 = 0.05$ to $\tau_1 = 0.0$ represents a Pareto improvement based on the unconditional utility function, but is no longer a Pareto improvement in terms of the conditional utility function. If the society values no addiction as a virtue, since $\tau_0 = 0.05$ is associated with lower stock of addictive good than $\tau_1 = 0.0$, the positive tax rate may be socially more desirable. We now use the simulation results presented in Table 2 to illustrate that the SOF-based evaluation can overcome this conflict. With $\alpha = 0$, evaluation by the SOF must satisfy the weak Pareto principle. Therefore, the SOF value should be higher for $\tau = 0$ than for $\tau = 0.05$. For small enough values of α, the SOF value needs to be higher for $\tau = 0$. From Table 4, we observe that for $\alpha \leq 0.05$, we get higher SOF values when $\tau = 0$ than when $\tau = 0.05$. However, for large enough values of α, the SOF value can be smaller for $\tau = 0$ than the one associated with $\tau = 0.05$. For instance, for $\alpha = 0.05$ the SOF is maximized at $\tau = 0.05$ which renders any other bequest tax rate (including $\tau = 0$) undesirable. In this way, the evaluation based on the SOF can resolve the limitation of the Pareto principle in an economic environment with endogenously determined preferences.

5. Tough Love Altruism with Bequest

Imagine a three-period model economy with three agents; the representative parent, the representative child, and the government. For simplicity, we consider the case of a single parent and a single child. The three periods considered are childhood, work and retirement for the child. We make the following seven assumptions. First, the timing of the model is assumed to be such that the life of the parent and the child overlap in the first two periods of the child's life. Hence, the parent has the child in the second period of his own life, which in turn corresponds to the first period of the child's life. Second, the parent not only cares about his own consumption, but is also altruistic toward the child. He assigns a weight of θ to the child's lifetime utility, where $0 \leq \theta \leq 1$.[14] Third, in period 2 of his life the parent receives an exogenous income, denoted by y^P. For simplicity, we assume that the parent receives no income in the last period of his life, but simply divide savings from the previous period into his own consumption and bequest. The bequest is taxed at the rate of τ by the government. Fourth, the parent maximizes utility over the last two periods of his life by choosing consumption, inter-vivos transfers, and bequest, denoted by C^P, T, and B, respectively. Fifth, the child is assumed to be a non-altruist, and derives utility only from her own consumption stream $\{C_t^K\}_{t=1}^3$.[15] y_2^K denotes child's second period exogenous income, and we assume that she receives no income in the first and last period of her life.

[14] When compared to the framework of Bhatt and Ogaki (2012), we have the following relationship:

$$\theta = \widetilde{\beta}\left(\frac{1-\eta}{\eta}\right)$$

[15] In this simple consumption good economy, we view consumption as a composite good that may include leisure activities such as TV time, video game time etc.

Sixth, the child's childhood consumption is assumed to be equal to the parent's inter-vivos transfers, because of social convention (alternatively, the child is assumed to be borrowing constrained in period 1 with a binding constraint) . Lastly, there is no uncertainty in the economy.

In the tough love model, the parent wants the child should grow to be patient, but is tempted to spoil the child. This interpretation is captured by the following two important features of the model. First, the child's discount factor is endogenously determined as a decreasing function of period 1 consumption:

$$\beta_K(C_1^K) \quad ; \quad \frac{d\beta_K}{dC_1^K} < 0.$$

We assume that the child's childhood consumption equals transfers from the parent ($C_1^K = T$). Therefore, the childs period t discount factor is given by $\beta_K(T)$. The idea is that if the child is spoiled by too much consumption during her childhood, then she will grow to be impatient.

Second, the parent does not use the child's endogenous discount factor, but uses a constant discount factor, $\beta_{t,P}$ to evaluate the child's lifetime utility. The parent's objective function is given by,

$$U_P(x) = u(C_2^P) + \widetilde{\beta} u(C_3^P) + \theta \Big(u(C_1^K) + \beta_P u(C_2^K) + \beta_P^2 u(C_3^K) \Big). \tag{17}$$

where $\widetilde{\beta}$ is the parent's own consumption discount factor and β_P is the discount factor used to evaluate the child's future utility, and θ denotes the altruism parameter.

The government collects the bequest tax from the parent, and distributes s as a lump sum subsidy. We assume that $s = \tau B$. An allocation in this economy consists of $x = (C_2^P, C_3^P, C_1^K, C_2^K, C_3^K)'$. The parent solves the following optimization problem:

$$\max_{C_2^P, T, B} \Big[u(C_2^P) + \widetilde{\beta} v(R(y_2^P - C_2^P - T) - B)\Big]$$

$$+\theta\Big[u(T) + \beta_P u(C_2^{K*}) + \beta_P^2 u(R(y_2^K + (1-\tau)B + s - C_2^{K*}))\Big], \tag{18}$$

subject to:

$$\{C_2^{K*}\} \equiv \arg\max_{C_2^K} \Big[u(C_2^K) + \beta_K(T) u(R(y_2^K + (1-\tau)B + s - C_2^K))\Big]. \tag{19}$$

where R is the gross interest rate, which is assumed to be exogenously fixed by a linear technology. In the above framework, the government can influence the child's patience by changing the bequest tax rate. If the bequest tax rate is reduced, then the parent has a greater incentive to leave bequests than to make transfers to the child. Lower transfers in turn would imply a higher discount factor for the child.[16]

5.1. Limitation of Pareto Efficiency

In this section we illustrate that the Pareto efficiency criterion may have difficulty in ranking social states when preferences are endogenously determined. For this purpose, we first define the child's unconditional utility function, which represents her *unconditional preference*

[16]It should be noted that the government's objective when setting the bequest tax rate may not have anything to do with affecting the child's preferences, but any nonzero tax rate does in fact affect her preferences.

ordering as follows:

$$U_K(x) = u(C_1^K) + \beta_K(C_1^K)u(C_2^K) + \beta_k(C_1^K)^2 u(C_3^K). \tag{20}$$

Given the state variable of the parent's transfer, T, the child's conditional utility function, which represents *conditional preference ordering*, is given by the following expression:

$$U_K(x|T) = u(C_1^K) + \beta_K(T)u(C_2^K) + \beta_k(T)^2 u(C_3^K). \tag{21}$$

We numerically solve the parent's optimization as a non-linear root finding problem. For the purpose of simulations, we assume the following functional forms for the period utility and the child's discount function:

$$u(C) = \frac{C^{1-\sigma}}{1-\sigma}. \tag{22}$$

The discount factor is given by:

$$\beta_K(T) = \beta_0 + \frac{1}{1+aT} \quad \textit{where} \quad a > 0 \textit{ and } \beta_0 \leq 0. \tag{23}$$

In our solution algorithm we impose the government's budget constraint: $s = \tau B$.[17]

Now, imagine that $\tau_0 = -0.15$ is the original policy situation. The government has been promoting bequests using this negative bequest tax rate. Consider a policy change to

[17]The details of our solution algorithm are provided in the appendix B.

eliminate this negative tax by setting the tax rate to zero: $\tau_1 = 0$. Let $x(\tau_i)$ be the allocation under the bequest tax rate of τ_i, $U_P(x(\tau_i))$ be the parent's utility under τ_i, and $U_K(x(\tau_i))$ be the child's unconditional utility under τ_i. Let $U_K(x|T(\tau_0))$ be the child's conditional utility given $T(\tau_0)$ (the equilibrium transfer when the tax rate is τ_0). The conditional utility is the child's retrospective evaluation of her lifetime consumption stream in the allocation x based on the grown-up child's utility function under the original policy regime.

Table 3 presents the optimal values of the unconditional utility for both child and parent, the discount factor of the child, and the conditional utility of the child. These are reported for both values of the bequest tax rate and utilize a given set of values for the model parameters.[18] As observed from this table, in terms of the unconditional utility function, we have a Pareto improvement when the policy changes from τ_0 to τ_1:

$$
\begin{aligned}
(24) \qquad & U_P(x(\tau_1)) > U_P(x(\tau_0)) \\
& U_K(x(\tau_1)) > U_K(x(\tau_0))
\end{aligned}
$$

The parent gains utility from the policy change because he gets more utility from succumbing to temptation to spoil the child. If the child is asked about the policy change during childhood, she will prefer being spoiled under the zero tax rate. However, in terms of the child's conditional utility function given the original tax rate, the child is made worse off by this policy change:

[18] We use the same parametric values as used by Bhatt and Ogaki (2012).

Table 3: Pareto Efficiency and Policy Evaluation: Tough Love Altruism

Global Parameters

$\theta = 0.51$; $R = 0.4$; $\sigma = 1.2$; $\beta_0 = -0.5$

$\widetilde{\beta} = \beta_p = 0.99$; $y_2^K = 1$; $y^P = 10$; $a = 0.18$

	$\tau_0 = -0.15$	$\tau_1 = 0.0$
$U_P(x^P(\tau_i))$	-16.8126	-16.8067
$U_K(x^K(\tau_i))$	**-6.8551**	-6.8241
$\beta_K(T(\tau_i))$	0.3107	0.3066
$V_K(x^K(\tau_1)\|T(\tau_0))$	-	**-6.8604**

(25) $$U_K(x(\tau_1)|T(\tau_0)) < U_K(x(\tau_0)|T(\tau_0))$$

If the child, after growing up to be patient under the negative tax policy, is asked in retrospect about the policy change, then she will prefer the negative tax rate. A society that values patience as a virtue may deem $\tau_0 = -0.15$ to be the socially more desirable policy when compared to the Pareto improvement (in terms of unconditional utility) represented by $\tau_1 = 0.0$.

5.2. Introducing Moral Virtues in Policy Evaluation

We now introduce moral virtues in the tough love altruism model and derive policy implications of such an extension. For this purpose we need to define the three evaluation fucntions, namely, SWF, MEF, and SOF. The SWF is defined as follows:

(26) $$SWF = U_p + U_k$$

where U_P and U_K are given by equations (17) and (20), respectively.

The MEF is given by:

$$MEF = -(\beta_K(T) - 1)^2 \tag{27}$$

so that larger deviations from the virtue of patience are morally undesirable.

An important component of the above formulation of the MEF is the definition of the virtue of patience. We define perfect patience as the time discount factor being exactly one.[19]

In the context of intertemporal choice models, Bhatt (2014) discusses the arguments for and against the view that zero discounting is a moral virtue. He argues that the common arguments against zero discounting conflate the normative with the positive aspects of the debate. Bhatt (2014) identifies two common criticisms against the view that zero discounting is a moral virtue. First is a lack of empirical evidence for such discounting behavior, and second is the undesirable implications of zero discounting for optimum consumption path in certain economic environments (Koopmans (1967), Olson and Bailey (1981)). He argues that although both are important elements in understanding individual choice, they do not serve as a normative basis for discounting. He finds that the ethical foundation for zero

[19] An important point here is to distinguish between intragenerational discounting and intergenerational discounting. Our definition of the virtue of patience concerns with the intragenerational discounting where we seek the normative value for the discount factor for discounting future utilities over one's own lifetime. On the other hand, intergenerational discounting concerns the discounting of the well-being of future generations. The issue of intergenerational discounting and the implied social discount rate is a key parameter in public policy debate. For instance, see the climate change debate surrounding the Stern Review (Stern, 2007). Some economists have criticized the social discount rate value used by the report as being too low (Nordhaus (2007), Weitzman (2007), Dasgupta (2007)). However, even among these critics most are sympathetic to the view that from a normative perspective, the pure time preference rate should be zero (Cowen and Parfit (1992), Broome (1994), and Dasgupta (2007))

discounting as a moral virtue is fairly robust. Such a view is also supported by others in the field of economics and philosophy (Brink (2010), Broome (1994), Ramsey (1928)). In this paper, we employ the MEF to express a moral judgment that one has a duty to value one's future self exactly as much as one's present self. It is important to note that the dictate of our MEF formulation is normative and not prescriptive. When a child cultivates preferences such that she is *pleased* with this duty, she is said to have the moral virtue of patience. Observe that this sense of duty is expressed in terms of preferences in our model, rather than in terms of *actions*; the choice of how much to save depends on the interest rate even when one has the virtue of patience.[20] emphasizes, we need to model the decision-making process when the sense of duty expressed by the MEF affects individual behaviors. For example, one can model the voting behavior of the child in the model when she feels that the MEF expresses her sense of duty and when she is tempted to vote for more spoiling. That type of modeling is beyond the scope of this paper.

For the purpose of defining the SOF we have to account for the fact that MEF and SWF are in different units and hence not directly comparable. Following the approach of section 4.2, we first define the two functions for the worst case scenario:

$$\overline{SWF} = U_p(x_0) + U_K(x_0) \tag{28}$$

$$\overline{MEF} = -\,(\beta_K(T_0) - 1)^2 \tag{29}$$

In the above definition of the $\overline{SWF}$, we utilize the worst possible allocation (x_0) in terms

[20] In order to model the free choice that George (1984)

of the SWF for the parent and the child.[21] We assume that the worst possible value for the moral evaluation function is obtained when the child receives the maximum possible transfers, because in that case her discount factor will be the lowest possible. In our model, $T_0 = y^P$ and hence we use $\overline{MEF} = -\left(\beta_K(y^P) - 1\right)^2$ in our simulations. The SOF is then given by the following expression:

$$SOF = (MEF - \overline{MEF})^{\alpha} \times (SWF - \overline{SWF})^{1-\alpha} \tag{30}$$

where $0 \leq \alpha \leq 1$ is the parameter of the SOF that sets the relative weights given to the moral virtue and welfare considerations.

5.2.1. Simulation Results

We solve the parents optimization problem numerically and use the same parametric specification and parameter values as in Section 5.1, for a menu of bequest tax rates. We assume that the tax rates available to the government range from -0.5 to 0.5, with an increment of 0.05. Table 4 presents the resulting optimal (i.e., SOF-maximizing) bequest tax policies. The optimized values for the SOF are presented in the bold in the table.

We discuss simulations for four policy scenarios, each of which is consistent with one of four alternative principles guiding government policy. The first is based on laissez-faire, wherein the government avoids affecting preferences through policy action. In this case the

[21] In our simulations we assume that the minimum level of each agent's consumption is 0.001, and use this level for each agent's consumption in x_0.

government would set the tax rate to zero. The second is based on welfarism, which involves maximizing the social welfare (i.e, maximizing the $SOF(\alpha = 0)$). The third is based on our proposed framework that weighs both welfarism and moral virtue considerations in policy evaluation. This can be achieved by setting $\alpha \in (0, 1)$ and then by maximizing the social objective function (SOF). Finally, the fourth is based solely on moral virtue ethics and aims to maximize only the moral evaluation function (MEF). This obtained by setting $\alpha = 1$ in our model.

There are several findings of interest from the simulation results presented in Table 4. First, a policy based on laissez-faire may lead to a social cost in terms of lower welfare. This can be observed from the simulations corresponding to $\alpha = 0$ in 4. We observe that based on laissez-faire, the tax policy of $\tau = 0$ does not maximize the $SOF(\alpha = 0)$ and hence is not a welfare maximizing policy.

Second, if we follow the principle of welfarism, which seeks to only maximize social welfare ($SOF(\alpha = 0)$), the optimal tax policy is $\tau = 0.2$. Hence, the government can achieve a higher level of welfare in our model economy by abandoning laissez-faire and following welfarism. An important point to note here is that in this case the government policy is impacting the preferences of the child, leading to a lower level of patience.

Third, given that the government policy is affecting preferences when it follows welfarism, it seems irresponsible for the government to completely ignore the moral virtue consideration by setting $\alpha = 0$. A more balanced approach would be to assign positive weights to both the SWF and the MEF. As we observe from Table 4, for small values of $\alpha = 0.01$ the optimum bequest tax based on maximizing the SOF leads to a smaller but still positive tax rate. On the other hand, if the government chooses to put a larger weight on moral virtue ethics then the optimum tax rate becomes negative. For instance, with $\alpha = 0.1$ the optimal bequest tax rate is -0.35. An interesting policy scenario is that of setting $\alpha = 0.05$. In this case the SOF is maximized at $\tau = 0$. Thus in our model economy, a balanced consideration of both moral virtue ethics and welfarism can lead to a zero tax rate; this is superficially similar to laissez-faire, but the motivations for the policy recommendation are very different.

Fourth, an extreme case is when the government only pursues moral virtue ethics and sets $\alpha = 1$. We observe that even in this case, the optimum tax policy of $\tau = -0.5$ fails to perfectly attain the virtue of patience, because the corresponding level of $\beta_K < 1$.

Finally, because the economy in Table 4 is the same as the one in Table 3, it is of interest to investigate whether policy evaluation based on the SOF can resolve the limitation of the Pareto principle highlighted earlier. As we observed from the Table 3 simulations, the change in the bequest rate from $\tau_0 = -0.15$ to $\tau_1 = 0.0$ represents a Pareto improvement based on the unconditional utility function of the child, but is no longer a Pareto improvement in terms of the conditional utility functions of the child. If the society values patience as a virtue, since $\tau_0 = -0.15$ is associated with greater patience than $\tau_1 = 0.0$, the negative bequest tax rate may be socially more desirable.

We use the simulation results presented in Table 4 to illustrate that the SOF-based evaluation can overcome this conflict. With $\alpha = 0$, evaluation by the SOF must satisfy the weak Pareto principle. Therefore, the SOF value should be higher for $\tau = 0$ than for

$\tau = -0.15$. For small enough values of α, the SOF value needs to be higher for $\tau = 0$. From Table 4, we observe that for $\alpha \leq 0.05$, we get higher SOF values when $\tau = 0$ than when $\tau = -0.15$. However, for large enough values of α, the SOF value can be smaller for $\tau = 0$ than the one associated with $\tau = -0.15$. For instance, for $\alpha = 0.075$ the SOF is maximized at $\tau = -0.15$ which renders any other bequest tax rate (including $\tau = 0$) undesirable. In this way, the evaluation based on the SOF can resolve the limitation of the Pareto principle in an economic environment with endogenously determined preferences.

Table 4: SOF vs SWF: Tough Love Altruism

Global Parameters

$\theta = 0.51$; $R = 0.4$; $\sigma = 1.2$; $\beta_0 = -0.5$; $\widetilde{\beta} = \beta_p = 0.99$
$y_2^K = 1$; $y^P = 10$; $a = 0.18$

τ	-0.5	-0.35	-0.15	0	0.15	0.2
β_K	0.3195	0.3158	0.3107	0.3066	0.3024	0.3010
$SOF(\alpha = 0)$	80.7976	80.8560	80.9228	80.9597	80.9785	**80.9790**
$SOF(\alpha = 0.01)$	77.1939	77.2446	77.3012	77.3309	**77.3431**	77.3417
$SOF(\alpha = 0.05)$	64.3164	64.3413	64.3645	**64.3706**	64.3620	64.3546
$SOF(\alpha = 0.075)$	57.3831	57.3956	**57.4029**	57.3980	57.3799	57.3698
$SOF(\alpha = 0.1)$	51.1971	**51.1998**	51.1943	51.1807	51.1552	51.1430
$SOF(\alpha = 1)$	**0.8431**	0.8380	0.8310	0.8254	0.8195	0.8176

6. Conclusion

In this paper, we propose a new way to evaluate social states in models with endogenous preferences. In our approach, virtue ethics is used in combination with welfarism to evaluate policy alternatives. Using two example models of endogenous preferences, namely, rational addiction and tough love altruism, we illustrate a key difficulty with using the Pareto principle; a policy that gives a Pareto improvement in terms of an agent's unconditional preference ordering may not be deemed socially desirable based on the conditional preference ordering. Furthermore, in the rational addiction example we show that even with a small weight given to the moral virtue considerations in policy evaluation process, the optimal tax on the addictive good may be positive even when there is no external cost associated with such a behavior. This is in stark contrast to the existing literature that focuses either on externalities or hyperbolic discounting to rationalize such a tax. Finally, based on informal discussions, we believe that many economists object to the use of moral virtue ethics considerations in public policy evaluation, because such an approach involves the government influencing people's preferences. Using the tough love altruism model we illustrate that a government policy based solely on welfarism can also influence an agent's preferences in economic environments where preferences are endogenously determined. Hence, the government may be influencing people's preferences even when it does not use any moral virtue consideration in the policy evaluation process. On the other hand, when we place a certain weight on moral virtue ethics, we find that the optimum bequest tax rate can be zero. Thus, incorporating moral virtue ethics may result in a policy that does not affect people's preferences. This illustrates that introduction of moral virtue ethics need not automatically lead to greater governmental influence on people's preferences.

Given these findings, one important implication of our theoretical analysis is that whether or not a certain government policy *does* influence people's preferences is an empirical issue that is independent of whether or not we think that the government *should* influence preferences. We believe that an important direction for future research in public policy is to gather empirical evidence for or against models with endogenous preferences. For the tough love altruism model used in this paper, there already exists some empirical evidence. A starting point of any model with endogenous time discounting is that genetic factors do not completely determine time discounting. Using a unique data set of twins in Japan, Hirata et al. (2010) found empirical evidence in favor of this. Kubota et al. (2013a,b) found empirical evidence consistent with the tough love altruism model, using unique survey data for the U.S. and Japan. Similarly, Akkemik et al. (2013) found evidence that empirically supports the main predictions of the tough love altruism model using data from Germany, Turkey, and Turkish migrants in Germany. We believe that more efforts to empirically validate models of endogenous preferences are needed in order to provide better insights into the effect of government policies on preferences. In our view, robust empirical evidence on this issue will significantly inform the discussion on public policy evaluation.

References

Akkemik, A. K., M. Bulut, M. Dittrich, K. Göksal, K. Leipold and M. Ogaki, "Worldviews and Intergenerational Altruism - A Comparison of Turkey and Germany," Unpublished, 2013.

Becker, G. and K. M. Murphy, "A Theory of Rational Addiction," *The Journal of Political Economy* 96 (1988), 675–700.

Benjamin, D. J., O. Heffetz, M. S. Kimball and N. Szembrot, "Beyong Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Prference," *American Economic Review* (2014), forthcoming.

Bhatt, V., "No Discounting as a Moral Virtue in Intertemporal Choice Models," *Keio-IES Discussion Paper Series. DP2014-003* (2014).

Bhatt, V. and M. Ogaki, "Tough Love and Intergenerational Altruism," *International Economic Review* 53 (2012), 791–814.

———, "Rational Addiction and Optimal Taxation: A Reexamination Based on the Social Objective Function," Unpublished, 2014.

Bhatt, V., M. Ogaki and Y. Yaguchi, "Normative Behavioral Economics Based on Unconditional Love and Moral Virtue," *Japanese Economic Review* 66 (2015), 226–246.

Bisin, A. and T. Verdier, "The Economics of Cultural Transmission and Socialization," in A. B. Jess Benhabib and M. O. Jackson, eds., *Handbook of Social Economics*volume 1A (The Netherlands: North-Holland: Elsevier, 2011), 339–416.

Brink, D. ., "Prospects for Temporal Neutrality," in C. Callender, ed., *Oxford Handbook of Time* (Oxford: Clarendon Press, 2010).

Broome, J., "Discounting the Future," *Philosophy & Public Affairs* 23 (1994), 128–156.

Bruni, L. and R. Sugden, "Market Reasoning as Moral Reasoning: Why Economists

Should Re-engage with Political Philosophy," *Journal of Economic Perspectives* 27 (2013), 141–164.

Constantinides, G. M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *The Journal of Political Economy* 98 (1990), 519–543.

Cowen, T. and D. Parfit, "Against the Social iscount Rate," in P. Laslett and J. Fishkin, eds., *Philosophy, Politics, and Society* (New Haven, NY: Yale University Press, 1992).

Cunha, F. and J. J. Heckman, "The Technology of Skill Formation," *American Economic Review* 97 (2007), 31–47.

Dahlsgaard, K., C. Peterson and M. E. P. Seligman, "Shared Virtue: The Convergence of Valued Human Strengths Across Culture and History," *Review of General Psychology* 9 (2005), 203–213.

Dasgupta, P., "National Institute Economic Review," *Philosophy & Public Affairs* 199 (2007), 4–7.

Doepke, M. and F. Zilibotti, "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," NBER Working Paper No. 20214, 2014.

Frey, B. S., *Happiness: A Revolution in Economics* (Cambridge, MA: MIT Press, 2008).

George, D., "Meta-Preferences: Reconsidering Contemporary Notions of Free Choice," *International Journal of Social Economics* 11 (1984), 92–107.

Goldfrab, R. and W. Griffith, "Amending the Economists Rational Egoist Model to Include Moral Values and Norms Part 2: Alternative Solutions," in K. Koford and M. J, eds., *Social Norms and Economic Institution* (Ann Arbor, MI: University of Michigan Press, 1991), 59–84.

Gruber, J. and B. Kőszegi, "Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes," *Journal of Public Economics* 88 (2004), 1959–1987.

Hausman, D. M. and M. S. McPherson, "Economics and Contemporary Moral Philosophy," *Journal of Economic Literature* 31 (1993), 671–731.

Heckman, J. J. and S. Mosso, "The Economics of Human Development and Social Mobility," *Annual Review of Economics* 6 (2014), 689–733.

Heckman, J. J., R. Pinto and P. A. Savelyev, "Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review* 103 (2013), 2052–86.

Hirata, K., H. Iiboshi, K. Hayakawa, S. Ikeda, Y. Tsutsui and F. Ohtake, "Genetic Inheritance of Time-Discounting Behavior: A Bayesian Approach Using Markov Chain Monte Carlo Method," Paper presented at the 2010 Japanese Economic Association Spring Meeting, Chiba University, Chiba, Japan, 2010.

Hirschman, A. O., "Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse," *Bulletin of the American Academy of Arts and Sciences* 37 (1984), 11–28.

Kaneko, M. and K. Nakamura, "The Nash Social Welfare Function," *Econometrica* 47 (1979), 423–435.

Kaplow, L. and S. Shavell, "Any Non-Welfarist Method of Policy Assessment Violates the Pareto Principle," *Journal of Political Economy* 109 (2001), 281–286.

Kőszegi, B. and M. Rabin, "A model of reference-dependent preferences," *The Quarterly Journal of Economics* 121 (2006), 1133–1165.

Koopmans, T., "Objectives, constrains, and outcomes in optimal growth models," *The Journal of Political Economy* 35 (1967), 1–5.

Kubota, K., A. Kamesaka, M. Ogaki and F. Ohtake, "Cultures, Worldviews, and Intergenerational Altruism," Paper presented at the European Regional Science Association 2013 Congress, Palermo, Italy, 2013a.

———, "Time Discounting and Intergenerational Altruism," Unpublished, 2013b.

Lawrence, C. J., M. Eichenbaum and C. L. Evans, "Nominal rigidities and the dynamic effects of a shock to monetary policy," *The Journal of Political Economy* 113 (2005), 1–45.

Mulligan, B. C., *Parental Priorities and Economic Inequality* (Chicago, IL: The University of Chicago Press, 1997).

Nordhaus, W. D., "A Review of the Stern Review on the Economics of Climate Change," *The Journal of Economic Literature* 45 (2007), 686–702.

Olson, M. and M. J. Bailey, “Positive Time Preference,” *The Journal of Political Economy* 89 (1981), 1–25.

Pollak, R. A., “Endogenous Tastes in Demand and Welfare Analysis,” *American Economic Review* 68 (1978), 374–379.

Ramsey, F. P., “A mathematical theory of saving,” *The Economic Journal* 38 (1928), 543–559.

Sandel, M. J., *Justice: What’s the Right Thing to Do?* (New York, NY: Farrar Straus & Giroux, 2009).

———, “Market Reasoning as Moral Reasoning: Why Economists Should Re-engage with Political Philosophy,” *Journal of Economic Perspectives* 27 (2013), 121–140.

Sen, A. K., “The Impossibility of a Paretian Liberal,” *Journal of Political Economy* 78 (1970), 152–157.

———, “Choice, Orderings and Morality,” in S. Korner, ed., *Practical Reason* (New Haven, NY: Yale University Press, 1974), 55–66.

———, “Rational Fools: A Critique of the Behavioral Foundations of Economic Theory,” *Philosophy and Public Affairs* 6 (1977), 317–344.

———, “Utilitarianism and Welfarism,” *The Journal of Philosophy* 76 (1979), 463–489.

Stern, N., *The Economics of Climate Change: The Stern Review* (Cambridge, UK: Cambridge University Press, 2007).

TEMKIN, L., *Rethinking the Good: Moral Ideas and the Nature of Practical Reasoning* (Oxford: Oxford University Press, 2011).

WEITZMAN, M. L., "A Review of the Stern Review on the Economics of Climate Change," *The Journal of Economic Literature* 45 (2007), 704–724.

A.1 Proof of Proposition 1

In this appendix we provide proof for *Proposition 1.* We begin by first deriving a general expression for the derivative of the SOF with respect to tau.

$$SOF = (MEF - \overline{MEF})^{\alpha}(SWF - \overline{SWF})^{(1-\alpha)} \tag{A.1}$$

Taking derivative with respect to τ, we get:

$$\frac{dSOF}{d\tau} = \left(\frac{MEF - \overline{MEF}}{SWF - \overline{SWF}}\right)^{\alpha}\left[\alpha\left(\frac{SWF - \overline{SWF}}{MEF - \overline{MEF}}\right)\frac{dMEF}{d\tau} + (1-\alpha)\frac{dSWF}{d\tau}\right] \tag{A.2}$$

Because $\left(\frac{MEF - \overline{MEF}}{SWF - \overline{SWF}}\right)^{\alpha} > 0$, we get:

$$sign\left(\frac{dSOF}{d\tau}\right) = sign\left[\alpha\left(\frac{SWF - \overline{SWF}}{MEF - \overline{MEF}}\right)\frac{dMEF}{d\tau} + (1-\alpha)\frac{dSWF}{d\tau}\right] \tag{A.3}$$

Hence,

$$\text{(A.4)} \qquad \left(\frac{dSOF}{d\tau}\right) > 0 \quad if \quad \frac{dMEF}{d\tau} > -\left(\frac{1-\alpha}{\alpha}\right)\left(\frac{MEF-\overline{MEF}}{SWF-\overline{SWF}}\right)\left|\frac{dSWF}{d\tau}\right|$$

Now, using the definitions of MEF and SWF, and the first order conditions for the decision-maker's maximization problem, it can be shown that equation (A-4) is satisfied if:

$$\text{(A.5)} \qquad a_0^*\frac{da_0^*}{d\tau} < 0.5\left(\frac{1-\alpha}{\alpha}\right)\left(\frac{MEF-\overline{MEF}}{SWF-\overline{SWF}}\right)\left|\tau\beta_1\frac{dq(C_1^*)}{dC_1^*}\left(R\frac{da_0^*}{d\tau}+\frac{da_1^*}{d\tau}\right)\right|$$

Given the conditions in *proposition 1*, the above inequality will be satisfied implying that the optimum tax rate is positive.

A.2 Proof of Proposition 2

B. Solution Algorithm

In this appendix we explain the numerical optimization method we used to solve the decision-maker's problem outlined in Section 3.2.

Step 1: Given T and B, the child solves the following optimization problem:

(B.1) $$\max_{C_2} \frac{C_2^{1-\sigma}}{1-\sigma} + \beta_k \frac{[R(y_2 + (1-\tau)B + z - C_2)]^{1-\sigma}}{1-\sigma}$$

where

$$\beta_k = \beta_0 + \frac{1}{1 + a(y_1 + T)}$$

The above optimization problem gives us a closed form solution for optimal values of C_2 and C_3:

(B.2) $$C_2^* = \frac{R(y_2 + (1-\tau)B + z)}{R + (\beta_k R)^{\frac{-1}{\sigma}}}$$

(B.3) $$C_3^* = R(y_2 + (1-\tau)B + z - C_2^*)$$

Step 2: We substitute for optimal C_2 and C_3 in the objective function and solve the parent's optimization problem:

(B.4) $$\max_{T,B} W \frac{[R(y_p - T) - B]^{1-\sigma}}{1-\sigma} + \theta\left(\frac{T^{1-\sigma}}{1-\sigma} + \beta_k \frac{C_2^{*1-\sigma}}{1-\sigma} + \beta_k^2 \frac{C_3^{*1-\sigma}}{1-\sigma}\right)$$

where

$$W = \frac{1 + \tilde{\beta}(\tilde{\beta}R)^{\frac{1-\sigma}{\sigma}}}{[R + (\tilde{\beta}R)^{\frac{1}{\sigma}}]^{1-\sigma}}$$

The step 2 optimization problem has no closed form solution for T and B. Hence we use numerical methods to find the solution to the above function. For this purpose we define a grid for T and B and choose a baseline for model parameters. Given these we search for the values of T and B that yields the maximum value for the objective function defined in equation (A-4). To implement this we need to initialize values of three key variables: T, B and the level of subsidy, i.e., z. For a given tax level set by policy, τ, we adopt the following algorithm to choose initial values:

1. For a given τ_i, we set:

$$T_{0i} = T^*(z^*_{i-1}; \tau_{i-1})$$

$$B_{0i} = B^*(z^*_{i-1}; \tau_{i-1})$$

2. For choosing the initial level of the subsidy we use:

$$z_{0i} = \tau_i B^*(z^*_{i-1}; \tau_{i-1})$$

We initialize the above process by first solving for the laissez-faire policy of $\tau = z = 0$.