POLITICAL ECONOMY RESEARCH INSTITUTE

PERI

University of Massachusetts Amherst

Discounting Nordhaus

Thomas R. Michl

January 2008

POLITICAL ECONOMY RESEARCH INSTITUTE

Gordon Hall
418 North Pleasant Street
Amherst, MA 01002

Phone: 413.545.6355
Fax: 413.577.0261
peri@econs.umass.edu
www.peri.umass.edu

UMASS

WORKINGPAPER SERIES

Number 158

Discounting Nordhaus

Thomas R. Michl
Colgate University
Department of Economics
Hamilton, New York, 13346
tmichl@mail.colgate.edu

January 9, 2008

Abstract

Discounting Nordhaus
by Thomas R. Michl
JEL Q5, E6
Keywords: Global warming, Stern Review, Discounting, Ramsey equation, Cambridge equation, Cambridge Theorem

This paper evaluates Nordhaus's neoclassical complaints about the *Stern Review* from the vantage point of classical growth theory. Nordhaus argues that the *Stern Review* exaggerates the effects of global warming because it uses a discount rate that is well below the market rate of return on capital. From the perspective of classical growth theory, Nordhaus's belief in choosing preference parameters for the social planner based on observed market rates of return filtered through the Ramsey equation is equivalent to assigning the preferences of the capitalist agents to the social planner. This equivalence is an implication of the Cambridge Theorem, which interprets the Ramsey equation as the saving function of the capitalist agents. The classical theory of growth interprets the market return to capital as a reflection of the property relations of capitalist society that does not offer the social planner any information that would be useful in resolving the problem of global warming. Contrary to the viewpoint of neoclassical economic theory, the market return to capital offers no information about preferences for the social welfare function or about the putative "marginal product" of conventional capital.

The purpose of studying economics is not to acquire a set of ready-made answers to economic questions, but to learn how to avoid being deceived by economists. —Joan Robinson

Nowhere is Joan Robinson's gnomic observation more applicable than in the debates about discounting spawned by the *Stern Review on the Economics of Climate Change* (2007). The back-to-back reviews of the *Review* by William Nordhaus (2007) and Martin Weitzman (2007) display some potential confusion about the discount rate created by modern neoclassical economics. Since both reviewers (misleadingly, we will argue) adopt the voice of a disinterested scientific observer, it is particularly important for non-economists and students of economics who want to engage with this public conversation to have an alternative account so that they can make up their minds informed by the full pluralist range of economic theory.[1]

This paper provides an alternative viewpoint that is grounded in the classical economic tradition of Adam Smith, David Ricardo, and Karl Marx as it has been translated through modern figures like Nicholas Kaldor, Joan Robinson, and Luigi Pasinetti. The classical growth model provides an alternative account of the return on capital that stands at the center of the discounting debate. The paper assumes some familiarity with the terms of modern economic discourse. While Weitzman agrees with most of the points that Nordhaus raises about discounting, he focusses on the economics of uncertainty. Therefore, we concentrate on Nordhaus's claim that the *Stern Review*'s high estimates of the cost of global warming depend on its choice of an indefensibly low discount rate.

1 The Ramsey and Cambridge equations

This debate centers around the Ramsey equation that connects growth to the rate of return on capital (often called the rate of interest by neoclassical economists and the rate of profit by classical economists). The Ramsey equation can be seen as the solution to an optimization problem by a so-called dynastic agent that lives forever. The idea of a dynastic agent was developed after Ramsey; he viewed his equation as the solution to an optimization

[1]For a broader discussion of the economics of global warming that is well informed by the criticisms of neoclassical economy theory raised by heterodox or non-orthodox economists, see Ackerman (2007).

problem by a benevolent social planner trying to achieve a social optimum for current and future generations. This is the standpoint of the *Stern Review* and its critics, who are asking important questions about how such a planner should behave, since this is thought to provide insight into what kind of policies real decision-makers ought to pursue. But it is best to start with the real economy and work back to the idealized planner.

Because discrete time is easier to visualize than continuous time, and in order to be able to make direct comparisons with the discrete-time classical growth models developed in Foley and Michl (1999) and Michl and Foley (2004), we will work with discrete periods rather than continuous time. The *Stern Review* and Nordhaus both operate in continuous time, and we provide some notes translating between the two frames of reference. In discrete time, C_0 refers to consumption in period $t = 0$. All variables are dated like this, but where the subscript refers to an arbitrary period, t, it is suppressed, and where it refers to the next period, $t + 1$, it is shortened to $+1$.

The dynastic agent that lives forever is a synthetic individual, created by observing that if each generation is altruistic in a particular way toward their children, the whole dynasty will behave as if it were one infinitely-lived dynastic individual. In this case we can write out its utility function for each generation like this:

$$U = \frac{C^{1-\alpha}}{1-\alpha}$$

This particular function is called a constant elasticity of substitution (CES) utility function because it can be demonstrated that the parameter α is the reciprocal of the elasticity of substitution between consumption (C) now and one period from now (C_{+1}). The *Stern Review* uses a function like this to represent the world's welfare, and sets α to unity. In this particular case, the function reduces to the simpler $U = \ln C$; this is sometimes described as the Cobb-Douglas form.[2] But α can take any value greater than or equal to zero. It is traditional to rule out a zero value based on the psychological law of diminishing marginal utility; $\alpha = 0$ implies that utility is a linear function of consumption so each additional unit of C produces one more util of pleasure. Larger values of α represent more sharply diminishing marginal utility.

[2]The Cobb-Douglas production function has an elasticity of substitution between its two inputs that is equal to one.

The dynasty's problem is to maximize this function over time, for the current and future generations. Because each cohort or generation is altruistic toward its children, and because people have a psychological propensity to prefer present consumption to future consumption, the dynasty discounts (shrinks down) future consumption by a factor, β.[3] Different authors use different notation for this parameter, the pure time preference factor, and its continuous time counterpart, the pure rate of time preference. The dynasty chooses a consumption sequence that maximizes[4]

$$(1-\beta)\sum_{t=0}^{\infty}\beta^t U(C)$$

The $(1-\beta)$ term has been placed outside the summation for aesthetic reasons; it simplifies the solution by letting β serve as both the pure time discount factor and, as we will see below, the saving propensity out of wealth.

Pure time discounting shows up inside the summation. Utility generated farther in the future counts less because it is discounted by a smaller fractional factor as t rises. For example, if $\beta = .95$, one util of pleasure generated by consumption 100 years in the future will be worth only .0003 utils today.[5]

To anticipate, in evaluating the effects of global warming the *Stern Review* uses a discount factor much closer to unity (equivalently, a discount rate near zero), and this is Nordhaus's complaint:

> In fact, using the *Review*'s methodology, more than half of the estimated damages 'now and forever' occur after the year 2800. The damage puzzle is resolved. The large damages from global warming reflect large and speculative damages in the far-distant

[3] If people prefer future consumption over present consumption (e.g., because a vacation is more fun if you get to anticipate the experience), β might be greater than unity, but almost all discussions presume that people are impatient.

[4] For comparison with Nordhaus's paper, note that in continuous time this would be an integration problem, and the pure rate of time preference would be treated like an interest rate in a present discounting formula:

$$\int_{t=0}^{\infty} U(C)\exp((\beta-1)t)$$

Here the discount factor, β, has been changed to a discount rate, $(1-\beta)$ that goes to zero (no discounting) when $\beta = 1$.

[5] That is $(1-.95)(.95)^{100}$.

future magnified into a large current value by a near-zero time discount rate (Nordhaus, 2007, p. 696).

Since we are constructing a classical growth model, let us assume that this dynasty is a class of capitalist agents surviving solely from its wealth or capital, K, that receives a rate of return, r. What the capitalist agent saves and invests becomes its wealth in the next period, K_{+1}, so its budget constraint is simply $K_{+1} + C = (1 + r)K$. Both its wealth and consumption grow by the factor $(1+g) = K_{+1}/K = C_{+1}/C$ in a steady state. Given a rate of return, the solution to this optimizing problem is the Ramsey equation in discrete time:[6]

$$(1 + g) = (\beta(1 + r))^{(1/\alpha)}$$

The Ramsey equation lies as the center of the controversies about discounting, and we will return to it repeatedly.

1.1 Cobb Douglas utility and the Cambridge equation

In the special case where in case $\alpha = 1$, the Ramsey equation simplifies to a linear function:

$$(1 + g) = \beta(1 + r)$$

Here capitalists are saving a constant fraction of their end-of-period wealth (i.e., including the profits it generates). This has been a staple assumption of classical growth theorizing for over two centuries. A version of it is sometimes called the Cambridge equation, because it is used by modern classical writers associated with Cambridge University, such as Kaldor or Robinson. The *Stern Review*'s choice of Cobb-Douglas utility fortuitously aligns with this essay's theme.

To visualize the underlying economics in this case, we can easily derive the approximation that $r - g \approx 1 - \beta$. This illustrates that the amount that each generation saves and passes along to the future depends on how much

[6] You can solve this problem using the Lagrangian method outlined in Foley and Michl (1999, ch. 5), together with the method of "guess and verify." See the appendix for the solution. For comparison with Nordhaus's review, the Ramsey equation in continuous time is $g = (r - (1 - \beta))/\alpha$, where the discount factor, β, has been replaced by a discount rate, $(1 - \beta)$, that is deployed like an interest rate used to calculate present value.

the future is discounted by pure time preference. Discounting penalizes the dynasty for consuming in the future. In the extreme case where the future is not discounted at all ($\beta = 1$), all profits are saved and the rate of profit and the rate of growth will be equal because there is no penalty for consuming in the future.

1.2 CES utility

If $\alpha \neq 1$, the growth rate will be a non-linear, but still increasing, function of the rate of profit. Most attention has been devoted to the case of $\alpha > 1$, in which case growth will be lower than in the case $\alpha = 1$. The economic intuition is that now there is an additional penalty to passing along wealth to the future: future cohorts will have more wealth and consumption, and with such a large α they will experience sharply diminishing marginal utility. There is an advantage here to "redistributing" consumption toward the present generation because it is poorer than future generations. The term "discounting" thus can refer to a pure rate of time preference or to the penalty from diminishing marginal utility.

This property of α becomes significant when the CES utility function is used as a social welfare function, because it measures the payoff from redistributing income progressively. Large values of α imply a large payoff from reducing the consumption of the rich in order to raise the consumption of the poor.

2 Classical growth models

The simplest way to build a classical growth model is to include the dynastic capitalist agent in a model with workers who do no saving at all and just live hand-to-mouth. This is not as restrictive as it might seem. A more realistic and satisfying assumption about worker saving would be the life-cycle hypothesis that workers save for their retirement. In this case, they are not altruistic toward their children, and leave no bequests, so the class structure reproduces itself over time. At any point in time, some workers will be saving but others will be dissaving in retirement, and at the social level it is clear that the net amount of worker saving depends on the ratio of retirees (dissavers) to workers (savers). This ratio is basically the growth of the labor force.

Without worker saving, it is clear that the Ramsey equation fully characterizes the relationship between the rates of profit and growth. But which determines which? The traditional classical answer (i.e., going back to Ricardo or Marx) is that the rate of profit is determined by the distribution of income and the technology, and both are exogenous to the accumulation decision, at least up to a first approximation. The distribution of income condenses down to the relation between the real wage and the productivity of labor, and this is a remarkably stable structural feature of capitalist society. The profit share, π, and its complement, the wage share $1 - \pi$, show considerable stability, both across time and across countries. The relevant technological variables are the ratio of output to capital stock (sometimes called the productivity of capital), ρ, and the depreciation rate of capital, δ. These also have been fairly stable. The rate of profit is by definition:

$$r = \pi\rho - \delta$$

Classical growth models that begin with this assumption are endogenous growth models because the rate of growth is a free variable that depends on capitalist saving behavior, technology, and the real wage. The underlying assumption is that labor supply does not constrain growth because capitalist economies are fundamentally labor-surplus systems. For example, there may be demographic or structural features that release increased labor supplies in response to the demands for workers generated by growth and accumulation. In this case, the Ramsey equation reads causally from right to left: the rate of profit determines the rate of growth.

At the other extreme, some modern classical growth models adopt the standard neoclassical assumption that the labor force limits growth, which in the long run at least takes place under conditions of full employment. (There might be some unemployment, but if there is an equilibrium unemployment rate, such as the natural rate or NAIRU theories predict, it amounts to the same thing.) This assumption makes the rate of growth fully exogenous, and the task of a growth model is to show how the distributional variables adjust to whatever it happens to be. In this case, the Ramsey equation needs to be rearranged to showcase the causal structure: the rate of growth determines the rate of profit. Specializing the utility function to the Cobb-Douglas form, the Cambridge equation written in its traditional form reads causally from right to left:

$$(1 + r) = \frac{(1 + g)}{\beta}$$

You might think that things would be more complicated if we introduced worker saving for life-cycle reasons into this model. Wouldn't worker saving make the Ramsey equation insufficient as a description of the relationship between profitability and growth?

2.1 The Cambridge Theorem

It turns out that the answer to that question is, surprisingly, no, at least in the long run. This was discovered by Luigi Pasinetti (1962) in the context of an exogenous growth model with $\alpha = 1$. He called this result the Cambridge Theorem, and it says that given a rate of growth, the rate of profit is determined by the Cambridge equation, independently of worker saving or technology. It seems reasonable to extend this theorem to the more general Ramsey equation.

The Cambridge Theorem applies to a two-class model (i.e., with worker saving included) of endogenous growth as well; see Michl and Foley (2004) for an example. In this case, it states that the rate of growth is determined by capitalist saving behavior, independently of worker saving, given the rate of profit.

In both these cases, the Cambridge Theorem describes the very long-run, steady state equilibrium. Take the case of an endogenous growth model, for example. An increase in worker saving in an endogenous growth model will *temporarily* increase capital accumulation. But because workers have a lower propensity to save than capitalists, the growth rate that would prevail if they were the only agents in the model is lower than the growth rate generated by capitalist saving (i.e., described by the Ramsey equation). It follows that eventually the growth of capital owned by workers will return to the growth rate established through the Ramsey equation. An increase in worker saving will have an effect on the distribution of capital wealth (workers will own a greater share), but not on its long-run growth rate.

Similar reasoning shows that in an exogenous growth model, an increase in worker saving will have no effect on the rate of profit in the long run, although it will lead to a higher share of capital being owned by workers. In both the exogenous and endogenous growth models, the Cambridge Theorem establishes that capitalists occupy a privileged position at the commanding heights in the structure of accumulation.

2.2 Interpreting the data

But what does all this have to do with global warming and discounting? We can observe the rate of profit and the rate of growth in real economies such as the US. The rate of return on capital lies somewhere in the range 7–15 percent per year, with 10 percent per year a fairly good estimate. The rate of growth of capital ranges from 2.5–5 percent per year. Thus, a rough estimate of 0.93–0.95 for β makes sense if the utility function is Cobb-Douglas. If it is more generally CES, then a range of (β, α) is possible. We can, in other words, indirectly observe the preferences of capitalist agents, assuming they are behaving according to our model.

One complication is that other models could produce the same behavior, characterized by the Ramsey or Cambridge equations. For example, if capitalist agents save for the sake of accumulating capital itself, either because they want the "warm glow" of giving it to their heirs or because they want the "warm glow" that comes from building up a fortune, that can lead to the same Cambridge equation written above. (I provide some examples in a forthcoming book (Michl, 2008) using a Cobb-Douglas utility function with consumption and capital as arguments. In this case, the Cambridge equation tells us about the weight given to capital accumulated at the end of life relative to consumption during life.)

If we want to recover the preference parameters of the worker agents, we could use the observed distribution of wealth to generate estimates. In our classical two-class model, their discount factor for future generations would be zero since the life-cycle theory assumes no role for intergenerational altruism.[7]

3 The social planner's problem

All this matters because Nordhaus and others would like us to believe that the Ramsey equation should be used by policy makers for guidance in selecting a discount factor for measuring the costs of global warming. Here is where we need to pay attention in order to avoid being deceived by economists.

The standard approach by economists to policy making in a growth setting is to ask how a benevolent social planner with absolute powers (i.e., a

[7]Translated into continuous time terminology, their discount rate for future generations would be infinity.

benevolent dictator) would allocate resources over time. One answer[8] assigns the planner the task of maximizing the discounted utility of all current and future generations, making no effort to distinguish between capitalist and worker agents. Letting the lower-case c and u represent consumption and utility for a homogeneous generation of people, the social welfare function might be:

$$(1-\beta_p)\sum_{t=0}^{\infty}\beta_p^t u(c)$$

where β_p represents the pure rate of time preference used by the planner. If we further assume that $u(c)$ takes the CES form, with parameter α_p, we can see that this problem shares some structure with the capitalist agent's problem, and the Ramsey equation will appear in some form in its solution.

From the perspective of the classical growth models surveyed earlier, it is clear that the social planner's problem only makes sense in the exogenous growth model because it requires that the population whose welfare is being optimized be well-defined. In the endogenous growth models, the population is itself endogenously determined by past accumulation, and is not well-defined in this sense, since a different history of accumulation implies a different population.

What separates the *Stern Review* from Nordhaus and its other critics concerns what values to assign β_p and α_p. The *Stern Review* takes α_p to be unity; it works with the simpler Cobb-Douglas form of the utility function. Its argument for this seems more pragmatic than principled.

On the other hand, the *Stern Review* takes the principled stance that β_p should be set (almost) equal to unity, implying virtually no discounting of future generations at all. This stance reflects a straightforward and powerful ethical position: people whose only crime is to be born in the future deserve to be treated equally with those of us who walk the Earth in the present.[9] It is important to notice that Frank Ramsey (1928) himself advocated precisely this position.

[8] Nordhaus (2007) does provide a helpful discussion of other approaches, but he ultimately accepts this one.

[9] Why not exactly 1? The *Stern Review* assigns a very small probability to a global disaster like an asteroid that eliminates a large chunk of the human population, and uses a discount factor slightly less than unity to accommodate this contingency. Note that to accommodate a discount factor exactly 1 would require that we drop the $(1-\beta_p)$ term, and it would create problems with infinities such as a utility sum that does not converge.

Nordhaus (2007, p. 691) takes the position that the *Stern Review* is being elitist in these choices:

> The *Review* takes the lofty vantage point of the world social planner, perhaps stoking the dying embers of the British Empire, in determining the way the world should combat the dangers of global warming. The world, according to Government House utilitarianism, should use the combination of time discounting and consumption elasticity the Review's authors find persuasive from their ethical vantage point.

Fair enough. But what is Nordhaus's ethical vantage point?

> I have always found the government House approach misleading in the context of global warming ... Instead, I would interpret the baseline trajectory, from a conceptual point of view, as one that represents the outcome of market and policy factors as they currently exist ... This approach does not make a case for the social desirability of the distribution of incomes over space or time of existing conditions, any more than a marine biologist makes a moral judgment on the equity of the eating habits of marine organisms ... (Nordhaus, 2007, pp. 691-92)

So Nordhaus asks us to regard him as a neutral observer, like a biologist. His preferred strategy is to extract estimates of the key parameters that match observations of real economies:

> ... in calibrating a growth model, the time discount rate and the consumption elasticity [i.e., α_p] cannot be chosen independently if the model is designed to match observable real interest rates and saving rates. To match a real interest rate of, say, 4 percent and a growth in per capita consumption of 1.3 percent per year requires some combination of high time discounting and high consumption elasticity.(Nordhaus, 2007, p. 694)

To illustrate the importance of matching real economies, Nordhaus provides a series of calibrated simulations from his DICE-2007 model of global warming and capital accumulation. These simulations generate an optimal carbon tax that measures the social cost of carbon. The social cost of carbon is the present value of the reduction in utility in the future caused by emissions. Using the contested alternative discount rates, the simulations reveal that the social cost of carbon reported by the *Stern Review* is an order of magnitude higher than the DICE baseline. But choosing a baseline recalibrated to $\beta_p = 1$ and α_p chosen to match the observed market return on capital generates a social cost of carbon that differs marginally from the DICE baseline. Nordhaus (2007, p. 700) concludes that "the central difference between the *Review* and many other economic models lies in the implicit real return on capital embedded in the model."

But now Nordhaus is asking us to endorse the outcomes observed in real economies. The social planner should adopt the preference parameters extracted from market outcomes. Nordhaus is interpreting the world through a particular theory of growth, the neoclassical theory, that suppresses or ignores the class structure of accumulation. The theory assumes that the Ramsey equation characterizes the behavior of one representative agent.

From the vantage point of the classical theory of growth, the Cambridge Theorem explains that those parameters reflect the preferences of the capitalist agents who occupy the commanding heights of the structure of accumulation. In practice, it is difficult to identify who the capitalist agents are in real economies, just as it is difficult to operationalize any economic category, even one as simple as "investment". But a reasonable estimate might identify the top 1 or 2 percent of households, ranked by wealth or income, with the capitalist agents whose bequest saving dominates the accumulation process in advanced capitalist economies.

Nordhaus's position can now be seen as an evasion of some very difficult political questions that simply adopts the preferences of the financial and economic elites and assigns them to the social planner. Because the capitalist agents discount their own future generations (either from a pure preference for present consumption, or from sharply diminishing marginal utility experienced by the richer cohorts of the future), so should policy makers who presumably are charged with protecting the future of humanity. Stated in this way, it is hard to see Nordhaus as a biologist studying the "eating habits of marine organisms."

We should not let Nordhaus confuse us with such phrases as "the *Review*'s radical revision of the economics of climate change." Like his rhetorical use of "House utilitarianism" in the quotations above, this tropism creates heat, not light. In fact, the *Stern Review* takes a defensible position that traces back to quite respectable origins; the eponymous Frank Ramsey is one of the most significant figures in economics and mathematics in the twentieth century.

4 Reevaluating the *Stern Review*

From the vantage point of the classical growth theory, the market rate of return does not offer any information that would be of much use to a social planner. The fact that the *Stern Review* ignores the market rate of return in choosing a calibration for the social planner appears fit and proper. In particular, it is hard to argue with Stern's principled objection to applying a pure rate of time preference that penalizes people for when they are born.

On the other hand, the *Stern Review*'s decision to use a Cobb-Douglas utility function appears somewhat arbitrary. However, here we can consult the two-class growth model described above because it tells us how the overwhelming majority of people behave. Cobb-Douglas utility in a life-cycle setting implies that workers are saving out of their wage income for retirement. A change in the return on saving–the rate of profit to a first approximation–will not affect their saving rate out of wage income. In economic terms, with a constant elasticity of substitution equal to unity, the substitution effect (higher profit rate makes future consumption cheaper) is exactly offset by a wealth effect (higher profit rate makes any saving worth more today). Many if not most studies of saving behavior have indeed found that saving rates are not particularly sensitive to real returns on saving (Bosworth and Burtless, 1992); this is a standard reason given in macroeconomics textbooks for leaving the rate of return out of the consumption function. It follows that the Cobb-Douglas form is a decent approximation to average behavior.

But should observed behavior be the benchmark for a social planner? We have already rejected the idea that it should in the case of a pure rate of time preference. One could, however, argue that the social planner would want to use observed market behavior to determine how much utility (whatever that is) is produced by the consumption of each generation. In this case, the *Stern Review*'s choice of a log utility function (with a near-one pure time discounting factor) makes good sense as an approximation to average behavior. Yet it is hard to avoid the feeling that this decision has been taken *faute de mieux*, and that it serves the role of a placeholder until something better comes along. Nordhaus sensibly complains that the *Stern Review* fails to conduct a robustness analysis using a range of calibrations. Future generations will probably be richer than we are, and perhaps more capable of absorbing the costs of global warming.

5 Optimal growth for workers

Economists have studied the question of what distribution of income would be optimal for workers who save for life-cycle purposes. Given a rate of growth (i.e., in an exogenous growth model), the optimal allocation of consumption over each typical worker's life span will arise when the rate of profit equals the rate of growth. That is called the *golden rule*. From the Cambridge equation, $(1+g) = \beta(1+r)$ we can see that this requires a capitalist discount factor of unity. Through this perspective, the gap between the rate of profit and the rate of growth measures how far market economies are from an allocation that is optimal from the point of view of workers. If the social planner could force the capitalist agents to abstain from consuming out of their wealth, each generation of worker would be as well-off as possible, given the constraint that society needs to maintain accumulation at a predetermined rate. This vantage point contextualizes Nordhaus's and others' frequent complaints that the *Stern Review* recommends an implausibly high saving rate in the present.

6 The marginal product of capital

There is one final, separate argument that Nordhaus makes about the market rate of return: it measures the "marginal product" of conventional capital. This is supposed to give the planner an indication of how much to invest in "climate capital" such as emissions abatement. If the hurdle rate of return to climate capital is set too low by the planner, it will result in an inefficient program of overinvestment in climate capital and underinvestment in conventional capital. Nordhaus (2007, p. 695) objects to the *Stern Review*'s use of a low real rate of interest for precisely this reason: "The efficient strategy has more investment in conventional capital at the beginning and can use those additional resources to invest heavily in climate capital later on."

This statement invites the question of the extent that it is predicated on the assumption of a well-behaved neoclassical production function in which more capital per worker has a well-defined, measurable relationship to output per worker. In particular, a marginal product from the addition of one unit of capital is well-defined. This theory has long been known to be theoretically unsupported; it lacks rigorous microeconomic foundations.[10] Economists understand so little about the relationship between capital accumulation and productivity that the only really defensible position for a social planner is one of relative ignorance. We know that the productivity of capital has remained constant at large time scales (with some fairly sizeable shifts, but no sustained trends) in advanced capitalist countries.[11] We also know that the productivity of labor has been increasing more or less steadily. We do not know how variations in capital accumulation affect those paths. A prudent planner would want to take the future course of technology as part of the given conditions by projecting those trends, or even better a range of trends. The idea that capital represents a technology for transferring resources from the present to the future, like planting a tree that bears fruit in the future, owes more to the quasi-theological discourse of marginalist economics than to any scientific principle.

The social planner is interested in the technological structure of the problem of global warming, including the impact of research and development on the productivities of labor and capital; the rate of profit does not in itself provide any additional information that would help her out. Recall that the rate of profit is the arithmetic product of the profit share and the productivity of capital. The market share of income going to the owners of capital in the form of profit plays no role in the planner's problem because it is not part of its "primitives" or fundamental conditions.

Even if we take the well-behaved production function to be merely a useful "parable" (a defense often offered by more sophisticated neoclassical theorists), the role of technical choice is arguably subsidiary to the distribution of income in a classical growth model. Once the rate of return on capital has been determined by the Cambridge equation, the production function is relevant only for determining which technique has a marginal product equal to the rate of profit. (This is why neoclassical economists sometimes call the Cambridge Theorem the "Pasinetti Paradox," since it effectively moots

[10]For elaboration of this point, see Cohen and Harcourt (2003).

[11]There have been many episodes of declining capital productivity, and these do create the impression of movement along a putative production function, which Foley and Michl (1999) and Michl (1999) call a fossil production function. However, the rate of profit almost universally exceeds the apparent marginal product of capital during these episodes, and does not convey any information about the technological frontier. See also Michl (2008).

the marginal productivity theory of distribution.) From the classical perspective, the debate over the production function remains secondary to the debate over the Ramsey/Cambridge equation rehearsed above.

7 Summing up

Why does all this matter? Nordhaus and other critics of the *Stern Review* advocate a climate-policy ramp that calls on policy makers to take modest steps toward emission reduction today, postponing the really heavy investment in abatement for the distant future. That *could* be (we are not trying to resolve this question here!) the wrong thing to do if the *Stern Review*'s estimates of the costs of global warming are accurate. We began with a quip from Joan Robinson, and we end with a modified version of another that is sometimes attributed to her: economics is too important to be left to the (neoclassical) economists.

8 Appendix: Deriving the Ramsey equation

Since I could not find any textbook derivation (or even presentation) of the Ramsey equation in discrete time using the CES form utility function, here is a derivation using the Lagrangian method that students and others might find useful.

The problem is to choose the sequence $\{C\}_0^\infty$ to

$$\max\,(1-\beta)\sum_{t=0}^{\infty}\beta^t C^{1-\alpha}/(1-\alpha)$$

$$\text{subject to } C + K_{+1} \leq (1+r)K$$

$$\text{given } K_0, r$$

The first line is the objective function. The second line is the budget constraint. Treated as an equality, as it should be since throwing away wealth is not rational, this immediately tells us something about the rate of growth and the rate of profit. Taking $t = 0$ as a representative period, we have

$$K_1 = K_0(1+g) = (1+r)K_0 - C_0$$

which implies that

$$C_0 = (r-g)K_0$$

On a steady state, where both C and K grow at the same constant rate, the consumption-wealth ratio will be the difference, $r - g$. We will use this insight below.

To solve this problem, first write the Lagrangian, which consists of the objective function minus the penalty function, with λ representing the penalty for violating the budget constraint, called the shadow price of capital by economists and the Lagrangian multiplier by mathematicians.

$$L = (1-\beta)\sum_{t=0}^{\infty}\beta^t C^{1-\alpha}/(1-\alpha) - \lambda(C + K_{+1} - (1+r)K)$$

Then obtain the first-order conditions that characterize a saddlepoint solution:

$$\partial L/\partial C = \frac{(1-\beta)\beta^t}{C^\alpha} - \lambda \leq 0 \tag{1}$$

$$\partial L/\partial K_{+1} = \lambda - \lambda_{+1}(1+r) \leq 0 \tag{2}$$

$$\partial L/\partial \lambda = C + K_{+1} - (1+r)K \geq 0 \tag{3}$$

From equation (1), evaluated at $t = 0$, we have:

$$C_0^{\alpha} = \frac{1-\beta}{\lambda_0}$$

Since we know that the penalty function will be zero from the saddlepoint property, we can write:

$$\sum_0^{\infty} \lambda C = \sum_0^{\infty} (\lambda - \lambda_{+1}(1+r))K_{+1} + \lambda_0(1+r)K_0$$

Now use equation (2), which will be satisfied as an equality because capital will never go to zero (it always provides some utility in future), to eliminate the first term on the RHS of this equation; and use equation (1) to replace the LHS with

$$\sum_0^{\infty} \lambda C = (1-\beta) \sum_0^{\infty} \beta^t C^{1-\alpha}$$

which gives us

$$\lambda_0 = \frac{(1-\beta)\sum_0^{\infty} \beta^t C^{1-\alpha}}{(1+r)K_0}$$

Substituting into the equation for C_0 above,

$$C_0^{\alpha} = \frac{(1+r)K_0}{\sum_0^{\infty} \beta^t C^{1-\alpha}}$$

The denominator of this expression can be expanded:

$$\sum_0^{\infty} \beta^t C^{1-\alpha} = C_0^{1-\alpha} + \sum_1^{\infty} \beta^t C^{1-\alpha}$$

Then the expression simplifies to

$$C_0 = (1+r)K_0 - C_0^{\alpha} \sum_1^{\infty} \beta^t C^{1-\alpha}$$

To resolve the summation on the RHS, we "guess" that the solution is a steady state, with C and K both growing at the same rate, g. Thus, $C_t = C_0(1+g)^t$ can be used to clean up the summation:

$$C_0 = (1+r)K_0 - C_0 \sum_1^{\infty} \beta^t (1+g)^{t(1-\alpha)}$$

Solving for C_0:

$$C_0 = \frac{(1+r)K_0}{1 + \sum_1^{\infty} \beta^t (1+g)^{t(1-\alpha)}} = (1 - \beta(1+g)^{1-\alpha})(1+r)K_0$$

This reads like a consumption function: consume a constant fraction of end-of-period wealth. Note that with Cobb-Douglas utility ($\alpha = 1$), that fraction does not depend on the rate of accumulation, and we can go back to the budget constraint to derive g immediately. Otherwise we use our guessed steady state, which implies as we saw above that $C_0 = (r-g)K_0$, to derive the Ramsey equation:

$$(1+g) = (\beta(1+r))^{1/\alpha}$$

Since every period is like the first, this equation can be generalized to describe the whole optimal program, verifying that our postulated solution solves the original problem.

References

Ackerman F., 2007. Economics for a warming world, *post-autistic economics review*, , no. 44, pp. 2–18

Bosworth B. and Burtless G., 1992. The effects of tax reform on labor supply, investment, and saving, *Journal of Economic Perspectives*, volume 6, no. 1, pp. 3–25

Cohen A. and Harcourt G., 2003. Retrospectives: Whatever happened to the Cambridge capital theory controversy?, *Journal of Economic Perspectives*, volume 17, no. 1, pp. 199–214

Foley D.K. and Michl T.R., 1999. *Growth and Distribution*, Cambridge MA: Harvard University Press

Michl T.R., 1999. Biased technical change and the aggregate production function, *International Review of Applied Economics*, volume 13, no. 2, pp. 193–206

Michl T.R., 2008. *Capitalists, Workers, and Fiscal Policy: A Classical Model of Growth and Distribution*, Cambridge MA: Harvard University Press

Michl T.R. and Foley D.K., 2004. Social security in a classical growth model, *Cambridge Journal of Economics*, volume 28, no. 1, pp. 1–20

Nordhaus W.D., 2007. A review of the *Stern Review on the Economics of Climate Change*, *Journal of Economic Literature*, volume 55, no. 3, pp. 686–702

Pasinetti L.L., 1962. Rate of profit and income distribution in relation to the rate of economic growth, *Review of Economic Studies*, volume 29, no. 4, pp. 267–279

Ramsey F., 1928. A mathematical theory of saving, *Economic Journal*, volume 38, pp. 543–559

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

Weitzman M.L., 2007. A review of the *Stern Review on the Economics of Climate Change*, *Journal of Economic Literature*, volume 55, no. 3, pp. 703–724