Federal Reserve Bank of New York
Staff Reports

The Paradox of Toil

Gauti Eggertsson

Staff Report no. 433
February 2010
Revised March 2010

The Paradox of Toil
Gauti Eggertsson
Federal Reserve Bank of New York Staff Reports, no. 433
February 2010; revised March 2010
JEL classification: E52, E62

Abstract

This paper proposes a new paradox: the paradox of toil. Suppose everyone wakes up one day and decides they want to work more. What happens to aggregate employment? This paper shows that, under certain conditions, aggregate employment falls; that is, there is less work in the aggregate because everyone wants to work more. The conditions for the paradox to apply are that the short-term nominal interest rate is zero and there are deflationary pressures and output contraction, much as during the Great Depression in the United States and, perhaps, the 2008 financial crisis in large parts of the world. The paradox of toil is tightly connected to the Keynesian idea of the paradox of thrift. Both are examples of a *fallacy of composition*.

Key words: zero bound, liquidity trap

Eggertsson: Federal Reserve Bank of New York (e-mail: gauti.eggertsson@ny.frb.org).

The author thanks Princeton University for hospitality during the writing of this paper and seminar participants at Princeton for comments. The author also thanks Michael Woodford for comments. The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Suppose you wake up one day and decide you want to work more. What happens? Presumably you start looking for some extra jobs, and if you are lucky enough you find some, perhaps in your own line of business, perhaps doing handy work, perhaps working in a bar in the evening. At the end of the day, holding everything else constant, you will work *at least* as much as you did yesterday and presumably more if you are lucky. If you think of society as a whole, even if you are only one member, your extra work generates some extra aggregate employment, some extra spending, and thus, perhaps – even if only a little – you may add to social output and welfare (whether your own welfare improves, of course, depends on your preference for leisure).

Now do the following thought experiment: Imagine that everyone wakes up that same day with exactly the same idea: to look for some more work. At first blush, the pedestrian logic sketched out above may seem to apply. Everyone would work just a little more – on average – and presumably in the process create more aggregate employment and output.

This paper is about a paradox: Under particular assumptions, and in a particular environment, if everyone tries to work more, this will in fact *reduce aggregate employment in equilibrium.* This is the paradox of toil. It corresponds to a classic fallacy of composition. Just looking at what one person does, holding *everything constant*, can be misleading once everything is aggregated. This is true even in a representative agent economy like the one studied in this paper, where all workers look and act exactly the same. The reason for this is *general equilibrium*: When everybody does something, *nothing is constant.* In models of general equilibrium, it can be highly misleading to build intuition and draw inferences looking at only one individual in *partial equilibrium.*

What peculiar environment and what strange conditions lead to the paradox of toil? The environment is a standard New Keynesian dynamic stochastic general equilibrium (DSGE) model that is widely used in policy-making institutions and is the backbone of publications in many leading academic journals. I subject this model to a simple shock – an intertemporal disturbance - which leads to zero nominal interest rates and downward pressure on inflation and output. This disturbance is a natural way to characterize the economic environment in the U.S. and other parts of the world following the crisis of 2008: that of zero interest rates and downward pressures on output and prices, as, e.g., argued in Krugman (1998), Eggertsson and Woodford (2003, 2004), Eggertsson (2009), and Christiano, Eichenbaum, and Rebelo (2009). The key question of this paper is: What if everybody tries to work more in response to this disturbance? This simple - and standard - environment, and the assumption of large enough intertemporal shock, gives rise to the paradox of toil.

Let's now think through the paradox of toil in a loose way, which will then be formalized in the model: Everybody trying to work more in response to the intertemporal disturbance puts downward pressure on current and future wages. What happens? Firms cut their prices today and in the future and stand ready to supply whatever is demanded at those prices. Then what? This leads to *expectations of deflation*, which increase the real interest rate - the difference between the nominal interest rate and expected inflation - and the central bank can't offset this by cutting the nominal rate because the rate cannot be below zero. Higher real interest rates lead to lower demand because people prefer to spend in the future rather than today, since prices are expected to be lower in the future (and thus the return on savings higher). Because of lower spending today, firms demand less labor. Thus, more labor supply leads to lower wages, more deflation, and higher real rates, which leads to less spending, which leads to less hiring of workers. Therein lays the paradox.

The paradox of toil is a close cousin to a classic paradox, the paradox of thrift. John Maynard Keynes (1933) asked the following question: What happens if an individual tries to save more? Holding everything else constant, that individual accumulates more savings. But what if everybody tries this at the same time? Their higher savings lead to less aggregate demand. The lower aggregate demand, in turn, leads to lower aggregate output and income, thus reducing the *ability of people to save*. In the end, then, aggregate savings decline. Thus, while the paradox of toil is that if everybody tries to work more there will be less aggregate employment, the paradox of thrift is that if everybody tries to save more there will be less aggregate savings. In the model of this paper, the paradox of thrift applies with equal force as the paradox of toil.[1]

There is a close parallel between these paradoxes. In most models in the long run both an increase in people's willingness to work and an increase in people's willingness to save increases output. Why? It shifts the production possibility frontier of the economy (by increasing steady-state capital and/or labor supply). Keynesian models, however, are about the short run, since, as Keynes famously declared, in the long run we are all dead. In the short run, in Keynesian models (new and old), it can sometimes happen that the economy is operating inside of its production possibility frontier. This can happen because of insufficient aggregate spending, i.e., too little demand. It is this very fact that gives rise to both paradoxes, although they show up in somewhat different forms.[2]

There are several implications of the paradox of toil for policy, at least if one takes the standard New Keynesian model seriously. Generally, the main implication of the paradox is that it is ineffective to design policies aimed at increasing aggregate supply in the short run when the problem is insufficient demand, i.e., insufficient spending on goods and services. Even worse, increasing aggregate supply in the short run can actually be counterproductive. More specifically, this suggests that policies that reduce marginal costs of firms may be counterproductive once their general equilibrium effect is taken properly into account in certain class of models. These policies may include temporarily cutting marginal labor taxes or reducing the minimum wage in response to the crisis. Both policies have been popular among economists and pundits in the U.S. during the current crisis. Perhaps more exotically, the paradox suggests that certain policies implemented by Franklin D. Roosevelt during the Great Depression, namely the National

[1] The paradox of thrift in this class of models is discussed in some detail in Christiano (2004) and Eggertsson (2009).

[2] The paradox of toil is mostly driven by the assumption of demand constrained economy which comes about due to the assumption of price rigidities. The reason why firms are not hiring more workers is that there is not enough demand for the firms' product. The reason is *not* high labor cost, or lack of willing hands to work.

Industrial Recovery Act (NIRA), which facilitated the monopoly power of firms and workers to prop up prices, may have been expansionary, contrary to a long-standing literature starting with Keynes (1933), Friedman and Schwartz (1963), and more recently Cole and Ohanian (2004) and Chari, Kehoe, and McGrattan (2006). For further discussion on this point, see Eggertsson (2008b), who uses the paradox of toil to argue that NIRA may have been expansionary.[3] Finally, the paradox suggests that other sources of lower marginal costs, such as a drop in oil prices, may be contractionary at zero interest rates. This may be of interest, since the onset of the current crisis was associated with, among other things, a collapse in the price of oil and other commodities, which people might have thought was helpful, as opposed to harmful as suggested by the model in this paper.

An alternative implication of the theory, of course, which some probably prefer rather than the exotic policy conclusions scetched out above, is that the paradox reveals some fundamental weaknesses in New Keynesian theory. While the standard New Keynesian model studied here is very specific in many respects, I conjecture that the same paradox occurs in a broad class of models where nominal spending determines aggregate output. If one wants to interpret the paradox from this perspective it poses a relatively broad challenge to this class of models.

Before going further, it is worth highlighting one property of the current analysis that is of some general interest. As already noted, under certain assumptions the increase in preference for toil can just as easily be interpreted as resulting from a reduction in marginal tax rates on labor. The finding of this paper suggests that while the "multiplier" of these tax cuts is positive when interest rates are positive, it flips signs and becomes negative at zero interest rates. Similar logic applies to most other fiscal instruments, i.e., their effectiveness is very different at zero interest rates than under normal circumstances, as discussed in Eggertsson (2009), who studies a range of other fiscal instruments and evaluates their quantitative significance. This suggests that a careless interpretation of the large number of empirical studies on the effectiveness of fiscal policy could be misleading when discussing policy today, since most of these studies rely only on data at positive interest rates.

2 The Model

Here I sketch out a New Keynesian model of the simplest kind (e.g., assuming a common labor market and no capital accumulation) omitting several details.[4] Households, which are of measure 1, live in standard utility functions

$$E_t \sum_{T=t}^{\infty} \beta^{T-t} \left[u(C_T) - \psi_T v(l_T) \right] \xi_T, \tag{1}$$

[3] See also Eggertsson and Woodford (2004) who analyze optimal VAT tax variations where the VAT tax shows up in a similar way as variations in preference for "toil" in this paper.

[4] Eggertsson (2009) presents a model with endogenous capital accumulation and shows that the same result applies. For a textbook treatment of this class of models see, e.g., Woodford (2003).

where β is a discount factor, C_T is a consumption aggregate, and l_T is labor supply. People like to consume, so the function $u(.)$ is increasing and concave but they don't like to work, so the function $v(.)$ is increasing and convex; both functions satisfy standard assumptions. There are two shocks to preferences: ξ_t affects how people evaluate utility today relative to in the future, while ψ_t affects people's relative valuation of consumption and disutility of working in a given period. The households' period budget constraint is

$$P_t C_t + B_t = (1 + i_{t-1})B_{t-1} + P_t W_t l_t + \int_0^1 \Pi_t(i)di - T_t, \tag{2}$$

where B_t is a one-period risk-free nominal bond, i_t is the one-period risk-free nominal interest rate, W_t is the real wage, and $\Pi_t(i)$ is profits from firm i distributed lump sum to the households. The household maximizes (1) subject to (2). Household optimal consumption decisions lead to the consumption Euler equation

$$u_c(C_t) = (1 + i_t)\beta E_t u_c(C_{t+1})\frac{\xi_{t+1}}{\xi_t}\frac{P_t}{P_{t+1}}, \tag{3}$$

while the households' optimal supply of labor gives rise to

$$W_t = \psi_t \frac{v_l(l_t)}{u_c(C_t)}. \tag{4}$$

Observe that the shock ξ_t only affects the price of consumption today relative to the future, while the shock ψ_t affects how the household evaluates the marginal utility of consumption relative to the marginal disutility of labor. The maximization plan of the household must also satisfy the transversality condition:

$$\lim_{T\to\infty} E_t \frac{B_T}{P_T} u_c(C_T) = 0. \tag{5}$$

There is a continuum of firms of measure 1 that are monopolistically competitive. We make this assumption, because we want to give them the power to select a price that may be different from that of their competitors without resulting in unbounded sales. Following the literature, we assume that the aggregate consumption of households, C_t, in fact refers to a Dixit-Stiglitz aggregate of a continuum of goods of measure 1 (each firm produces one good variety), $C_t \equiv \left[\int_0^1 c_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}}$ with an elasticity of substitution $\theta > 1$. Dixit and Stiglitz (1977) show that the optimal consumption decision of each household for good varieties then implies that each firm faces the demand function $c_t(i) = C_t(\frac{p_t(i)}{P_t})^{-\theta}$, where P_t is the Dixit-Stiglitz aggregate price index, $P_t \equiv \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}}$. Before going into details of the firm's problem, note that all production is consumed so that $c_t(i) = y_t(i)$ and

$$Y_t = C_t, \tag{6}$$

where we use the notation $y_t(i)$ for the firms' i production and Y_t for aggregate output.

Following Calvo (1983), let us suppose that once a firm selects its price, it stays constant for a stochastic number of periods.[5] More precisely, suppose each firm has an equal probability of reconsidering its price in each period. Let $0 < \alpha < 1$ be the fraction of firms with prices that remain unchanged in each period. For each firm that revises its prices in period t, the new price p_t^* will be the same. The problem that each firm faces at the time it revises its price is then to choose a price p_t^* to maximize

$$\max_{p_t^*} E_t \left\{ \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \lambda_T [(1-\tau) p_t^* Y_T (p_t^*/P_T)^{-\theta} - P_T W_T Y_T (p_t^*/P_T)^{-\theta}] \right\}.$$

Here I use λ_T to discount the profit flow of the firms. It is the Lagrangian multiplier on the households' budget constraint, i.e., it measures the marginal value of nominal income to the households. The term $(1-\tau)$ is a tax/subsidy introduced for analytical convenience, paid lump sum from and to households. The first-order condition of this problem is

$$\max_{p_t^*} E_t \left\{ \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \frac{u_c(C_T)}{P_T} (\frac{p_t^*}{P_T})^{-\theta-1} Y_T [(1-\tau) \frac{p_t^*}{P_T} - \frac{\theta}{\theta-1} W_T] \right\}. \tag{7}$$

Note that by the Calvo pricing assumption we can write the price index as

$$P_t = [(1-\alpha)(p_t^*)^{1-\theta} + \alpha P_{t-1}^{1-\theta}]^{\frac{1}{1-\theta}}. \tag{8}$$

Let us also define price dispersion as $\Delta_t \equiv \int_0^1 (\frac{p_t(i)}{P_t})^{-\theta}$. We can then relate aggregate labor and aggregate output by

$$l_t = \int_0^1 y_t(i) di = \int_0^1 Y_t (\frac{p_t(i)}{P_t})^{-\theta} di = Y_t \Delta_t. \tag{9}$$

Price dispersion, using the assumption of staggered prices, follows the law of motion:

$$\Delta_t = (1-\alpha)(\frac{p_t^*}{P_t})^{-\theta} + \alpha \Delta_{t-1}. \tag{10}$$

Without going into details about how the central bank implements a desired path for nominal interest rates, we assume that it cannot be negative so that

$$i_t \geq 0. \tag{11}$$

To close the model, we assume that the central bank follows a Taylor rule, whose parameters are detailed further below, in accordance with

$$i_t = \max(0, \phi(\frac{P_t}{P_{t-1}}, Y_{t,} \xi_t, \psi_t)). \tag{12}$$

An equilibrium can now be defined as a collection of stochastic processes $\{Y_t, C_t, l_t, i_t, P_t, p_t^*, \Delta_t, B_t, T_t\}$ given a path for the exogenous variables $\{\xi_t, \psi_t\}$ and the initial condition P_{-1}, Δ_{-1} that satisfy

[5] This assumption is made not only for simplicity, but also because this has become the most common assumption in the literature (and has been subject to relatively extensive empirical testing, beginning with the work of Gali and Gertler (1999) and Sbordone (2002)). Moreover the resulting firms' Euler equation has been derived from relatively detailed microfoundations, see e.g. Gertler and Leahy (2008), who derive it assuming physical menu costs, and Woodford (2009), who derives it assuming imperfect information.

equations (3)-(12). I assume fiscal policy is Ricardian, so the choice between debt and taxes is irrelevant. There are many different ways of writing down a fiscal rule that satisfies this property, and I omit further discussion of this issue (see, e.g., Woodford (2003) and Eggertsson and Woodford (2003) for a discussion).

Our analysis will be a local analysis around a steady state.[6] A steady state, i.e., a constant equilibrium without shocks, exists such that $\Delta_t = \frac{P_t}{P_{t-1}} = 1$, $Y_t = C_t = l_t = \bar{Y}$, and $i_t = \bar{r} = 1/\beta - 1$. Local to this steady state, we do a linear approximation of the model. First, by combining (7) and (10), it can be shown that price dispersion is only of second order, i.e.,[7]

$$\hat{\Delta}_t = 0,$$

implying that

$$\hat{Y}_t = \hat{l}_t. \tag{13}$$

The consumption Euler equation (combined with $Y_t = C_t$) is approximated to yield

$$\hat{Y}_t = E_t \hat{Y}_{t+1} - \sigma(i_t - E_t \pi_{t+1} - r_t^*), \tag{14}$$

where $\sigma \equiv -\frac{u_{cc}}{u_c \bar{Y}}$, $r_t^* \equiv \bar{r} + \hat{\xi}_t - E_t \hat{\xi}_{t+1}$, where the hat denotes deviation from steady state and $\pi_t \equiv \log P_t - \log P_{t-1}$ and $\bar{r} \equiv \log \beta^{-1}$. The nominal interest rate now denotes $log(1 + i_t)$ in terms of our previous notation so we can once again write the zero bound as

$$i_t \geq 0. \tag{15}$$

The consumption Euler equation says that aggregate demand (which is determined entirely by consumption demand) depends on future output and the real interest rate. The pricing equation of the firm can be approximated to yield

$$\pi_t = \delta \hat{W}_t + \beta E_t \pi_{t+1}, \tag{16}$$

where $\delta \equiv \frac{(1-\alpha)(1-\alpha\beta)}{\alpha}$. This equation says that a price the firm sets will depend on the real wage, which is the real cost of producing one extra unit of output for each firm. As this marginal cost increases, there is higher aggregate inflation according to this pricing relationship. Approximating equation (4), we obtain

$$\hat{W}_t = \omega \hat{l}_t + \sigma^{-1} \hat{Y}_t + \hat{\psi}_t, \tag{17}$$

where $\omega \equiv \frac{v_l l}{v_{ll}}$. Finally, the monetary policy rule is approximated by

$$i_t = \max(0, r_t^* + \phi_\pi \pi_t), \tag{18}$$

[6] For simplicity, I have set the parameter τ through lump-sum taxation so as to eliminate monopoly distortions in steady state. This does not affect the results in any way in the current exercise, but would be important if we wanted to explore welfare or Ramsey policies.

[7] Accordingly, I will omit further discussion of price dispersion in the paper and focus on the first-order approximation of the model.

where we assume that $\phi_\pi > 1$. This rule says that the central bank will adjust the nominal interest rate to offset the intertemporal shock. I do this to focus the discussion on the role of the shock to "toil", i.e. $\hat{\psi}_t$, without mixing it up with the policy reaction function.[8] If $\hat{\xi}_t$ was the only shock, this rule implements the first best allocation in the absence of the zero-bound and r_t^* corresponds to the "efficient" real interest rate. An approximate equilibrium can now be defined as a collection of stochastic processes $\{\pi_t, \hat{l}_t, \hat{W}_t, \hat{Y}_t, i_t\}$ that satisfy (13)-(18), given the exogenous processes $\{\hat{\psi}, \hat{\xi}_t\}$.

3 The long and the short run: equilibrium allocations

This section summarizes the model's equilibrium allocation in two propositions under a special assumption about the shocks. The rest of the paper is devoted to illustrating and interpreting the meaning of the equilibrium allocations shown in these propositions. Accordingly, the discussion of the propositions here is short and terse, with proofs relegated to the Appendix.

The *long run* in the model is defined as the time at which all shocks have gone back to steady state. The short run is the period in which the economy is subject to temporary disturbances. More precisely, the short run is defined by the particular values of the shocks $\hat{\xi}_S$ and/or $\hat{\psi}_S$. Both shocks revert back to steady state, 0, with probability μ in each period. They are perfectly correlated. Let us call the stochastic period in which the shocks are back to steady state τ. Then $t \geq \tau$ is defined as the long run, while $t < \tau$ is the short run. This assumption is very convenient because of the structure of the model. As we will see, it allows us to boil the model down to just a few static "short run" equations, even if in principle we are dealing with an infinite-horizon model. This draws out the paradox of toil in the most transparent way.

Proposition 1 In the long run, $t \geq \tau$, there is a unique bounded solution such that $\pi_t = \hat{Y}_t = \hat{W}_t = \hat{l}_t = 0$ and $i_t = \bar{r}$.

Proof: See Appendix.

We now state two conditions that are helpful to analyze the short run:

$$C1 \qquad r_S^* + \frac{\delta\phi_\pi}{1-\beta(1-\mu)+\frac{\sigma\kappa}{\mu}(\phi_\pi - 1 + \mu)}\psi_S < 0$$

$$C2 \qquad \mu(1-\beta(1-\mu)) - (1-\mu)\sigma\kappa > 0,$$

where $\kappa \equiv \delta(\omega + \sigma^{-1})$.

Proposition 2 In the short run, $t < \tau$, then we consider two cases:

[8] This makes no difference at the zero-bound and does not affect the paradox of toil. It does, however, have an effect at positive interest rate and I find this assumption to be the most useful for clarifying the logic of the model.

1.(*Positive interest rates in the short run*) If C1 does *not* hold, there is a unique bounded equilibrium at positive interest rates such that

$$\pi_t = \pi_S = \frac{\delta}{1-\beta(1-\mu)+\frac{\kappa\sigma}{\mu}(\phi_\pi - 1+\mu)}\hat{\psi}_S \tag{19}$$

$$\hat{W}_t = \hat{W}_S = \frac{1-\beta(1-\mu)}{1-\beta(1-\mu)+\frac{\kappa\sigma}{\mu}(\phi_\pi - 1+\mu)}\hat{\psi}_S \tag{20}$$

$$\hat{Y}_t = \hat{l}_t = \hat{Y}_S = \hat{l}_S = -\frac{\frac{\delta\sigma}{\mu}(\phi_\pi - 1+\mu)}{1-\beta(1-\mu)+\frac{\kappa\sigma}{\mu}(\phi_\pi - 1+\mu)}\hat{\psi}_S \tag{21}$$

$$i_t = i_S = r_S^* + \frac{\delta\phi_\pi}{1-\beta(1-\mu)+\frac{\sigma\kappa}{\mu}(\phi_\pi - 1+\mu)}\hat{\psi}_S. \tag{22}$$

2. (*Zero interest rates in the short run*) If C1 *and C2 hold,* there is a unique bounded equilibrium at zero interest rates such that

$$\pi_t = \pi_S = \frac{\kappa\sigma}{\mu(1-\beta(1-\mu))-\sigma\kappa(1-\mu)}r_S^* + \frac{\delta}{\mu(1-\beta(1-\mu))-\kappa\sigma(1-\mu)}\hat{\psi}_S \tag{23}$$

$$\hat{W}_t = \hat{W}_S = \frac{\kappa\sigma\delta^{-1}(1-\beta(1-\mu))}{\mu(1-\beta(1-\mu))-\sigma\kappa(1-\mu)}r_S^* + \frac{(1-\beta(1-\mu))}{\mu(1-\beta(1-\mu))-\kappa\sigma(1-\mu)}\hat{\psi}_S \tag{24}$$

$$\hat{Y}_t = \hat{l}_t = \hat{Y}_S = \hat{l}_S = \frac{(1-\beta(1-\mu))\sigma}{\mu(1-\beta(1-\mu))-\sigma\kappa(1-\mu)}r_S^* + \frac{\frac{(1-\mu)}{\mu}\sigma\delta}{\mu(1-\beta(1-\mu))-\kappa\sigma(1-\mu)}\hat{\psi}_S \tag{25}$$

$$i_S = 0. \tag{26}$$

Proof See Appendix.

We do not analyze the case in which C1 holds but C2 does not. In this case either the local approximation is no longer valid (the model explodes) or there is an indeterminacy of equilibria. Proposition 2 encompasses the core of the paper. In fact, the paradox of toil can be seen right away by noticing that while a negative $\hat{\psi}_S$ (i.e., people want to work more) increases output at positive interest rates according to equation (21), it reduces output according to equation (25) at zero interest rates. The rest of the paper analyzes the interpretation of this mathematical result and why the paradox of thrift springs to life at zero interest rates.

4 What happens if people want to work more?

Let us first consider a simple thought experiment proposed in the introduction. What happens if people want to work more? We first analyze this question under "normal circumstances," i.e., the zero bound is not binding, so condition C1 in the last section does not apply. Observe that ξ_S appears only in equation (14) so if $i_t > 0$, and we substitute the policy function (18) into (14), it drops out. Hence, the size of ξ_S is irrelevant if C1 does not apply since it will always be offset by movements in the nominal interest rate.

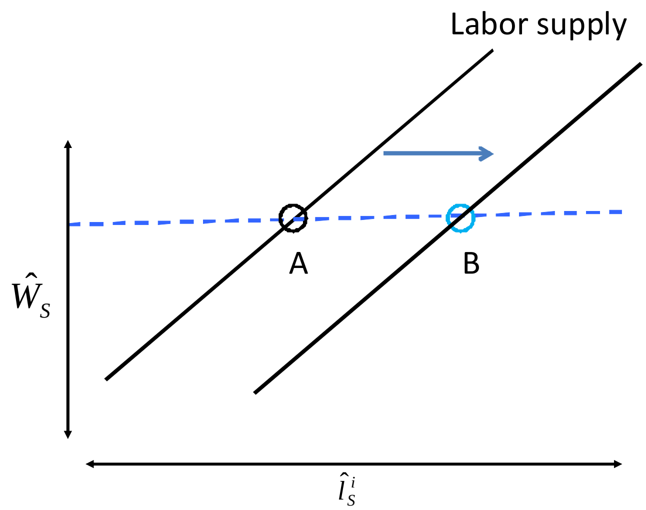

Figure 1: Partial equilibrium analysis of the effects of toil

Consider the following experiment: Suppose that people want to work more in the sense that $\psi_S < 0$ is temporarily lower at time t and then reverts back to steady state with probability μ in each period (the size of the ξ_S is irrelevant, as discussed above).

Recall from Proposition 1 that, in the long run, $t \geq \tau$, all variables are at their steady state, i.e., $\pi_t = \hat{Y}_t = \hat{l}_t = \hat{C}_t = 0$. Furthermore, in the short run, according to Proposition 2, all the variables take on a unique bounded constant value (since the shocks are constant in the short run). We denote all these constant short-run variables with the subscript S.

Let us first study partial equilibrium, i.e., what happens from the perspective of a single worker, i, who faces an exogenous wage $\hat{W}_S$. The labor supply is given by equation (17):

$$\hat{l}_S^i = \omega^{-1}\hat{W}_S + \omega^{-1}\sigma^{-1}\hat{C}_S^i - \omega^{-1}\hat{\psi}_S^i. \tag{27}$$

This equation is plotted in the (W_S, l_S^i) space in figure 1 - holding $\sigma^{-1}\hat{C}_S^i$ constant (which measures the marginal value of extra consumption) - and the curve is denoted by a dotted line. The curve is upward sloping in the (W_S, l_S^i) space: As the real wage rate increases, the household is willing to work more. What happens as the worker wants to work more, i.e., after a drop in $\hat{\psi}_t^i$ for a given $\sigma^{-1}C_S^i$? This curve shifts out. Because we are now analyzing this in partial equilibrium, i.e., holding marginal value of consumption $\sigma^{-1}\hat{C}_S^i$ and wages $\hat{W}_S$ constant, this will increase the work done by the worker by the line segment A-B. Hence, there are no surprises in this analysis, much as one would have expected from intuition. Now, let's go over the general equilibrium analysis, where we take into account both the change in marginal utility of consumption and the equilibrium effect on wages.

It is helpful to summarize the model in terms of two aggregate schedules: the equilibrium demand for labor services and the equilibrium supply of labor services. Let's start with the

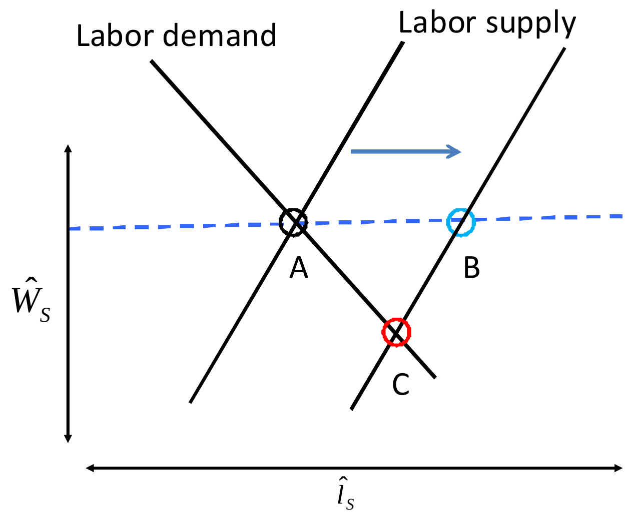

Figure 2: General equilibrium effect of an increase in toil at positive interest rates.

supply of labor. Again, consider equation (17). But now we use the equilibrium relationship that all output is consumed and that all income comes from aggregate labor, i.e., $\hat{C}_S = \hat{Y}_S = \hat{l}_S$. From this we obtain

$$LS \qquad \hat{l}_S^{Supply} = (\omega + \sigma^{-1})^{-1}\hat{W}_S - (\omega + \sigma^{-1})^{-1}\hat{\psi}_S. \tag{28}$$

This schedule is plotted in Figure 2. It shows the optimal supply of labor, for any given level of wages, which is determined by the households. Again, it is upward sloping in $\hat{W}_S$. But now we have taken into account a general equilibrium effect, i.e., the second term in (17). An increase in work leads to an increase in consumption, and this reduces the marginal utility of income, thus leading the worker to increase labor supply less for any wage increase. Hence, the aggregate labor supply curve - in general equilibrium - is steeper than in partial equilibrium. Again, a negative shock to ψ_S shifts the curve to the right, i.e., given any wage rate, the worker is then ready to work more. Consider first a "partial equilibrium" in the sense that the wage that workers face is fixed.[9] Then, once again, a shift in the labor supply leads to more work, corresponding to the segment A-B.

Our model, however, is a full general equilibrium model, and hence we need to understand how an increase in the aggregate supply of labor interacts with the demand for labor in general equilibrium. To do this, let us derive the aggregate demand for labor by the firms. In our model, the firms are monopolistic competitors: They set a price and then commit to supply whatever goods are demanded at that price. Think of a grocery store, for example, that fills its shelves with goods and posts a price. People can buy as much as they want of this good. Thus, to find

[9] This allocation would also apply if the central bank, in its reaction function, would fully accommodate $\hat{\psi}_S$ to offset its' effect on inflation, either by directly incorporate it into our definition of r_t^* or by $\phi_\pi \to \infty$.

the demand for labor at the store, we need to figure out the aggregate demand for the goods at the grocery store, which then pins down how much labor the firms wants to employ to sell the goods demanded. Consider first equation (14). Recall that in the periods $t \geq \tau$, the solution is $\pi_t = \hat{Y}_t = 0$. In periods $t < \tau$, then, inflation in the next period is either zero (with probability μ) or the same as at time t, i.e., $\pi_{t+1} = \pi_S$ (with probability $1-\mu$). Hence the solution in the short run, $t < \tau$, satisfies

$$\hat{Y}_S = (1-\mu)\hat{Y}_S - \sigma(i_S - (1-\mu)\pi_S - \bar{r}) = (1-\mu)\hat{Y}_S - \sigma(\phi_\pi + \mu - 1)\pi_S, \tag{29}$$

where I have substituted out for the interest rate policy rule, taking into account that the interest rate is positive according to Proposition 2, given that C1 does not apply. An important variable in determining how many goods consumers want to buy is inflation. From the pricing equation of the firm, we have

$$\pi_S = \delta\hat{W}_S + \beta(1-\mu)\pi_S.$$

Combining these two equations, using $\hat{Y}_S = \hat{l}_S$, we obtain the demand for labor:

$$LD \qquad \hat{l}_S^{Demand} = -\sigma\delta\mu^{-1}\frac{\phi_\pi + \mu - 1}{1-\beta(1-\mu)}\hat{W}_S. \tag{30}$$

We can see that this curve is downward sloping in the $(\hat{l}_{S,}\hat{W}_S)$ space. In other words, as the wage increases, the firms will demand less labor in equilibrium. What is the intuition for this? It can be seen in two steps. The first step is to observe that an increase in real wages increases inflation. An increase in the wage cost of the firm directly increases its marginal cost, i.e., it now costs more for the monopolist to produce each extra unit of the good it sells. As a result, firms that revisit their price (which they do at staggered intervals) will increase the price they charge for their goods. This, in turn, increases inflation in the aggregate, π_S. Now let's go to the second step in the logic. It is that inflation reduces demand when C1 does not apply so that interest rates are positive. To see how an increase in aggregate inflation maps into a reduction in labor demanded, we need to sort out how the increase in inflation changes the number of goods the household demands, because the firms are committed to supply whatever number of goods is demanded by the household at the prices they have set. Look at equation (29). Higher inflation leads the central bank to increase the nominal interest rate by the central bank by more than one-to-one ratio with inflation because $\phi_\pi > 1$. This increases the real rate, thus making output today more expensive than output in the future. Because of this change in the *price of output today relative to in the future,* people spend less. As people spend less, the firm produces less output, and thus there is less demand for labor. This completes the second step in the logic. To recap: Higher wages –>higher inflation. Higher inflation –> higher real interest rates –> less demand for goods by households –> firms hire less workers. Hence, labor demand is downward sloping in wages.

Having plotted both the demand and the supply for labor in Figure 2, we can now study the general equilibrium effect of an increase in the willingness of people to work more. An outward shift in the labor supply curve not only increases labor, but also reduces the real wage rate.

In general equilibrium, the fact that people want to work more creates more output and lower aggregate wages, forming the new equilibrium at point C. The exact analytical expression for point C can be seen from part 1 of Proposition 2. While this has been relatively standard analysis of the model, let's now turn to the interesting part.

5 The peculiar world of the zero bound

In this section, the short-term nominal interest rate is zero and there is a decline in output and prices, i.e., a recession. This environment creates the paradox of toil, the key focus of the paper. The main reason is that the demand for labor now becomes upward sloping in real wages. Consider now a shock such that C1 holds and the zero bound is binding. For now, let us set $\hat{\psi}_S = 0$. We will turn to movements in this shock in the next section. Here, we focus on the shock $r_S^* = \bar{r} + \hat{\xi}_t - E_t\hat{\xi}_{t+1} < 0$.

Since this shock, r_S^*, gives rise to the paradox of toil, it is perhaps appropriate to ask a basic question: Where does this shock come from? In the simplest version of the model, a negative r_S^* is equivalent to a preference shock as we have just seen and thus corresponds to a *lower* ξ_S in period $t < \tau$ that reverts back to steady state with probability $1 - \mu$ in each period. Everybody suddenly wants to save more, so the real interest rate must decline for output to stay constant. More sophisticated interpretations are possible, however. Curdia and Eggertsson (2010), building on Curdia and Woodford (2008), show that a model with financial frictions can also be reduced to the same set of equations analyzed in this paper. In this more sophisticated model, the shock r_t^* corresponds to an exogenous increase in the probability of default by borrowers. What is nice about this interpretation is that r_t^* can now be mapped into the wedge between a risk-free interest rate and an interest rate paid on risky loans. Both rates are observed in the data. This wedge exploded in the U.S. during the crisis of 2008, providing empirical evidence for a large negative shock r_t^*, and the same applies to the Great Depression. Del Negro, Eggertsson, Ferrero, and Kiyotaki (2009) tell a similar story in a bit more complex model. In their case, however, certain assets become "less liquid," but this also results in a collapse in the nominal interest rate and deflationary pressures. In any event, a financial crisis - e.g., characterized by an increase in probability of default by banks and borrowers - is my story for the model's recession.

Under our assumption, the shock r_t^* remains negative in the short-run S, until some stochastic date τ, when it returns to steady state. Then monetary policy takes the form:

$$i_t = \bar{r} \text{ for } t \geq \tau \tag{31}$$

$$i_t = 0 \text{ for } 0 < t < \tau. \tag{32}$$

Let's now revisit our demand and supply diagram for aggregate employment. The aggregate supply equation is unchanged, i.e., equation (28) is replicated here for convenience.

$$LS \qquad \hat{l}_S^{Supply} = (\omega + \sigma^{-1})^{-1}\hat{W}_S - (\omega + \sigma^{-1})^{-1}\hat{\psi}_S. \tag{33}$$

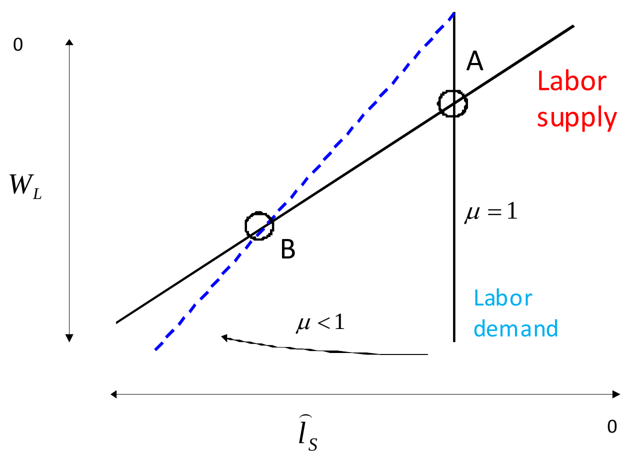

Figure 3: Equilibrium determination at zero interest rates

What changes is the *demand for labor*. Again recall that in the periods $t \geq \tau$, the solution is $\pi_t = \hat{l}_t = \hat{Y}_t = 0$. In periods $t < \tau$, inflation in the next period is either zero (with probability μ) or the same as at time t, i.e., $\pi_t = \pi_S$ (with probability μ). Consider first the consumption Euler equation:

$$\hat{Y}_S = (1-\mu)\hat{Y}_S + \sigma(1-\mu)\pi_S + \sigma r_S^*. \tag{34}$$

A key difference from before is that now the shock r_S^* appears because it can no longer be offset by interest rate cuts. The nominal interest rate is zero (hence, the term $\phi_\pi \pi_t$ drops out as opposed to the previous result). Now combine this equation with the pricing equation of the firms:

$$\pi_S = \delta \hat{W}_S + \beta(1-\mu)\pi_S \tag{35}$$

using $\hat{Y}_S = \hat{l}_S$ to yield the aggregate demand for labor

$$\hat{l}_S^{Demand} = \frac{\sigma(1-\mu)\delta\mu^{-1}}{1-\beta(1-\mu)}\hat{W}_S + \frac{\sigma}{\mu} r_S^*. \tag{36}$$

Again, it is helpful to graph the demand and supply equations in the $(\hat{l}_S, \hat{W}_S)$ space. Consider first a special case in which $\mu = 1$, i.e., the shock ξ_S reverts back to steady state in period 1 with probability 1. This case is shown in Figure 3. It applies only to the equilibrium determination in period 0. The equilibrium is shown where the two solid lines intersect at point A. From Proposition 2, we see that this equilibrium occurs below steady state, i.e., both output and employment are below their "first best" level. There is a recession. At point A, employment is now completely demand-determined by the vertical labor demand curve and pinned down by the shock r_t^*. For a given level of a labor demand, then, wages are determined by where the labor supply curve intersects the vertical labor demand. It is worth reemphasizing: *Employment is* *completely demand-determined.* It is also worth stressing that the firms here are monopolistic competitors and will thus supply whatever goods are demanded by customers given the prices that are posted. Hence, the demand for labor was derived by combining the pricing equation of the monopolistic competitor and the consumption Euler equation of the consumer, imposing $C_t = Y_t$. What really pins this equilibrium down is the number of goods the workers demand from equation (34), which again controls how much labor the firms hires to produce these goods.

Consider now the effect of $\mu < 1$. In this case, the contraction is expected to last for longer than one period. The expectation of a possible future contraction results in movement in the demand for labor so that the curve becomes flatter, i.e., the equilibrium is determined at point B. Observe that the demand-for-labor equation is no longer vertical but *upward sloping in wages*. The main reason for this comes from inflation, i.e., higher inflation *expectations* $(1-\mu)\pi_S$ increase output demanded according to equation (34). This is because for a given nominal interest rate ($i_S = 0$ in this equilibrium), any increase in expected inflation reduces the real interest rate, making current spending relatively cheaper than future spending and thus increasing demand. Conversely, expected deflation, a negative $(1-\mu)\pi_S$, causes current consumption to be relatively more expensive than future consumption, thus suppressing spending.

From equation (35) we see that there is a one-to-one relationship between real wages and inflation. Why is this so different from before? In the last section, the central bank reacted to the increase in inflation/deflation by increasing/decreasing the nominal interest rate by more than a one-to-one ratio. But now, inflation is below zero, i.e., below the implicit inflation target according to the policy rule (18) and the central bank cannot adjust the interest rate to a level below zero. Hence, the central bank is happy to see inflation rise (i.e., less deflation, to be precise) and will not increase the nominal interest rate as inflation (and expectations of future inflation) increases. This is the key to understanding why the slope of labor demand changes at zero interest rates and is now upward sloping.

Observe, furthermore, the presence of the expectation of future contraction, $(1-\mu)\hat{Y}_S$, on the right-hand side of equation (34). The expectation of future contraction makes the effect of both the short run shock (which is expected to persist) and the expected deflation even stronger through a negative $\hat{Y}_S$, which is why as μ declines (and the short run is expected to be "longer"), the fall in employment, wages, output, and inflation becomes much bigger. The fall in employment and wages, in fact, can go on without a bound as μ declines further until the model explodes. At a critical point $\bar{\mu}$, the two lines are parallel and no solution exists. It is easy to confirm from (33) and (36) that this occurs exactly when condition C2 is violated. Once we move beyond $\bar{\mu}$, the model has no unique bounded solution, and the equilibrium is indeterminate given our policy rule. One implication of the large effect of reducing μ, explored, for example, in Eggertsson (2009), is that the model can generate very large output and employment movements for reasonable parameters and relatively small shocks. This means, quantitatively, that problems such as the paradox of toil can be a big deal. And now it is time to turn to the paradox.

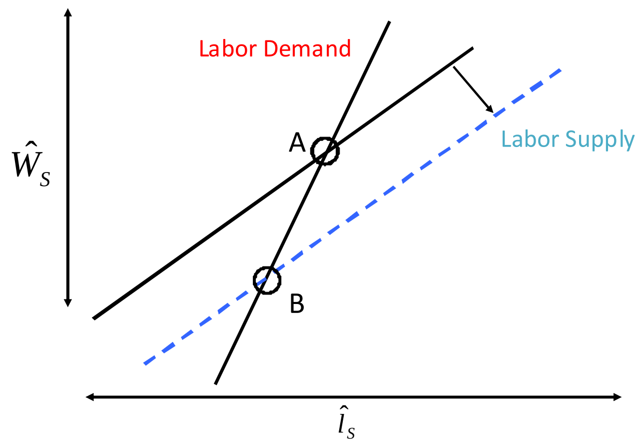

Figure 4: The paradox of toil

6 The paradox of toil

The paradox of toil appears right away in the plot derived in the last section. The change in the preference parameter ψ_s shifts labor supply. But now the labor demand is no longer downward sloping, it is *upward sloping in wages*! Furthermore, condition C2 guarantees that it is steeper than labor supply. Hence, the consequence of people wanting to work more is that, in equilibrium, everybody works less as can be seen in point B in figure 4. The logic is that the increase in labor supply lowers aggregate wages, which in turn gives the worker lower income to spend on goods and services. These deflationary pressures increase the real interest rate, and this cannot be offset by the central bank cutting the nominal interest rate. Thus less goods are demanded. Because the firms will employ labor only to satisfy any consumer demand for goods, this reduces aggregate employment. Why does the same logic not apply under normal circumstances? The reason is that, normally, when there is an outward shift in supply and downward pressures on wages and prices, the central bank will offset this by cutting of the nominal interest rate by more than one-to-one ratio because $\phi_\pi > 1$ in its policy rule. Hence, the real interest rate will decrease, making spending today cheaper and increasing output and employment. At zero interest rates, however, this is no longer possible due to the zero bound. Hence the paradox.

The intuition is further clarified in Figure 5, but some readers (the author included) find it easiest to develop intuition for results by looking directly at the dynamic equations. The negative shock ψ_t (people want to work more), and the expectation that it will persist, puts immediate deflationary pressure on prices in the AS equation (the right-hand-side equation in the figure) because it lowers the wage cost of firms. As segment A shows, this reduces inflation expectations in the AD equation (the left-hand-side equation), which increases the real rate of

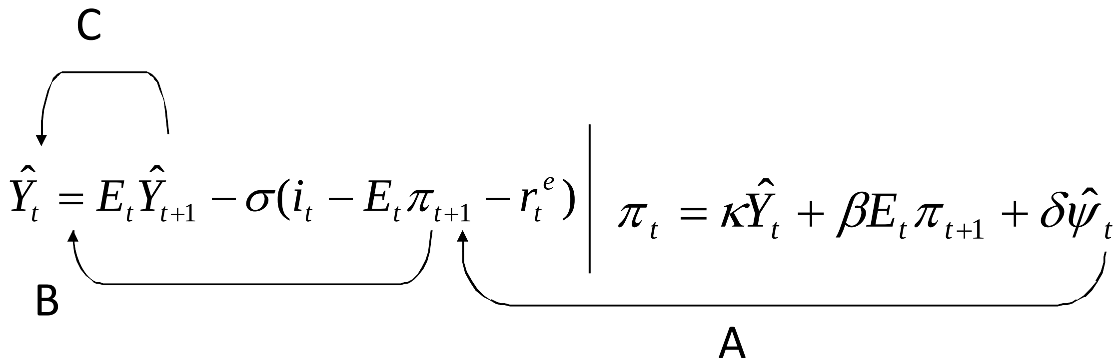

Figure 5: Some intuition

interest (because the interest rate is fixed at zero), thus suppressing demand, as shown by segment B. That's not the end of the story, however. People expect this to occur in the next period as well with some probability, leading to even further contraction denoted by segment C. This implies that the quantitative importance of the paradox of toil can be significant, as further discussed in Eggertsson (2009).

A short comment on the old Keynesian paradox of thrift. It also pops right out of this model. An increase in people's preference to save means a lower ξ_S, which makes r_S^* more negative. This shifts back the labor demand and leads to lower inflation and output. This paradox is more interesting once we have endogenous capital stock, since then one can show that if everyone tries to save more, aggregate savings collapses (in the current model, all output is consumed so there is no savings technology). This is seen in Eggertsson (2009), which also confirms the paradox of toil in a model with endogenous capital accumulation.

7 Implications for other experiments, such as tax cuts, the New Deal, and oil price hikes

The paradox of toil has a somewhat surprisingly implications for other policy experiments. Consider the first-order condition of the representative household reported above:

$$W_t = \psi_t \frac{v_l(l_t)}{u_c(C_t)}.$$

A labor tax that is proportional to wage income would enter in exactly the same way, i.e., $\psi_t = \frac{1}{1-\tau_t}$. The implication, of course, is that a temporary labor tax cut of this form *is contractionary according to the model.* It is worth stression, however, as further discussed in Eggertsson (2009) that while this specifaction of labor taxes is standar, it is special in many respect. Consider now the linearized pricing equation of the firm, where we have substituted the labor supply equation for marginal costs:

$$\pi_t = \kappa Y_s + \beta E_t \pi_{t+1} + \delta \hat{\psi}_t.$$

Multiple disturbances have been suggested that enter exactly the same way as $\hat{\psi}_t$ in this equation. Perhaps most obviously, variations in the parameter θ, which indexes the monopoly power of the firms, show up in this way, or variations in markups in the labor markets (e.g., an increase in the bargaining power of unions). This implies that a temporary increase in the monopoly power of firms (or unions) is expansionary in the model at zero interest rate, a point elaborated in more detail in Eggertsson (2008b). A similar comment applies to other increase in marginal costs, under certain assumptions, such as those due to increases in oil.

Before concluding, it is worth commenting briefly on one issue not yet discussed. We have assumed that labor markets are perfectly flexible; hence, wages adjust downward so that demand equals supply in the labor market at all times. This assumption was made mostly for the sake of exposition, because it draws out the paradox of toil most clearly. It is quite common, however, in many New Keynesian models to assume some rigidities in the wage-setting process. How might such rigidities change the results? Do they make the paradox worse? The answer is that rigidities in the wage setting make things a bit "better," because then an increase in labor supply does not exert as strong a downward pressure on the price level. We can think of rigidities in the wage setting as giving rise to a positive "labor wedge," in the terminology of Chari, Kehoe, and McGrattan (2006), which partially will offset the shock $\hat{\psi}_t$, which we can think of as a negative labor wedge.

One peculiar feature of the zero-bound environment, in fact, is that stronger rigidities in either prices or wages a*re stabilizing for the economy.* For prices, the simplest way to see this is to observe that an increase in the frequency of price changes in Figure 3 makes the labor demand curve flatter, thus moving point B further toward less employment and lower wages. The reason is that a higher degree of price flexibility means that people expect prices in the future to fall even further, thus raising deflationary expectations. This makes the real interest rate even higher, contracting demand even further. This possibility - that greater wage and price flexibility is destabilizing - was first suggested by Tobin (1975) and further analyzed in De Long and Summers (1986). Both papers argued that price flexibility can be destabilizing *in general,* while here we find this is only the case in the special environment created by the assumption of zero interest rates. This leads me to conjecture, that even if one considers price setting mechanism other than that of Calvo (1983), which implies more price flexibility, this could in fact amplify the paradox of toil.

So for we have assumed that the preference for toil moved one-to-one with ξ_t. But what if the change in ψ_t is permanent? In this case the result is no longer clear cut and depends on parameter values. Several forces are at work. A permenent increase in toil leads to a permament increase in output, which thus increases demand. Meanwhile, there can be a deflationary impact of a permanet increase of toil once the interest bound stops being binding, which works in the opposite direction. Whether the paradox applies under a permenent shif it toil, thus, becomes an empirical question.[10]

[10] The parameter configurations under which permanent increase in toil is expansionary can be seen from Proposition 5 in Eggertsson (2008b), but note that toil shows up in the same way as NIRA in that model. According to the

8 Conclusion

The main point in this paper is that increasing aggregate supply, e.g., desire to work more, is counterproductive at zero interest rate. This is the paradox of toil. Similar arguments can apply to other supply-driven shifts, under certain assumptions, such as cuts in marginal tax rates, lower oil prices, reduced monopoly power of firms or workers, and so on. A separate paper, Eggertsson (2009), shows, however, that policies aimed directly at stimulating aggregate spending work extraordinary well at zero interest rates. These policies include temporary sales-tax cuts, investment tax credits, and government spending.

9 References

Calvo, Guillermo (1983). "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics* 12, 383-398.

Chari, V. Kehoe, P., and McGrattan, E. (2006). "Business Cycle Accounting," Econometrica.

Christiano, Lawrence (2004). "The Zero-Bound, Low Inflation, and Output Collapse," mimeo, Northwestern University.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo (2009). "When Is the Government Spending Multiplier Large," Northerwestern, NBER Working Paper.

Cole, H., and L. Ohanian (2004). "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis," Journal of Political Economy, Volume 112, Number 4.

Curdia, Vasco, and Michael Woodford (2008). "Credit Frictions and Optimal Monetary Policy," mimeo, Columbia University.

Curdia, Vasco, and Gauti Eggertsson (2010). "What Caused the Great Depression?" mimeo, Federal Reserve Bankf of New York.

De Long, J.B., and Summers (1986). "Is Increased Price Flexibility Stabilizing?" American Economic Review 76, 1031-1044.

Del Negro, M., A. Ferrero, G. Eggertsson, and N. Kiyotaki (2010), "The Great Escape? A quantiative evaluation of the Fed's non-standard policy," mimeo, Federal Reserve Bank of New York.

Denes, Matthew, and Gauti Eggertsson (2009). "A Bayesian Approach to Estimating Tax and Spending Multpliers," New York Federal Reserve Staff Paper.

Dixit, A., and J. Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," American Economic Review, Vol. 67, No. 3, pp. 297-308.

Eggertsson, Gauti (2008a). "Great Expectations and the End of the Depression," *American Economic Review*, forthcoming.

Eggertsson, Gauti (2008b). "Was the New Deal Contractionary?" mimeo, New York Federal Reserve.

parameterization in that paper a permanent incraese in toil is contractionary under most parameter specification.

Eggertsson, Gauti (2009). "What Fiscal Policy Is Effective at Zero Interest Rates?, mimeo, New York Federal Reserve.

Eggertsson, Gauti, and Michael Woodford (2003). "The Zero Bound on Interest Rates and Optimal Monetary Policy," Brookings Papers on Economic Activity 1, 212-219.

Eggertsson, Gauti, and Michael Woodford (2004). "Optimal Monetary and Fiscal Policy in a Liquidity Trap," ISOM conference volume.

Friedman, Milton, and Anna Schwartz (1963). "A Monetary History of the United States," Princeton University Press.

Gali, Jordi, and Mark Gertler (1999). "Inflation Dynamics: A Structural Econometric Analysis," Journal of Monetary Economics 44: 195-222.

Gertler, Mark, and John Leahy (2008). "A Phillips Curve with an Ss Foundation," Journal of Political Economy. New York University.

Keynes, John Maynard (1933). "Open Letter to Franklin D. Roosevelt," New York Times, December 1933.

Keynes, John Maynard (1936). "The General Theory of Employment, Interest, and Money," London: Macmillan.

Krugman, Paul (1998). "It's Baaack! Japan's Slump and the Return of the Liquidity Trap," Brookings Papers on Economic Activity 2.

Sbordone, Argia (2002). "Prices and Unit Labor Costs: A New Test of Price Stickiness," Journal of Monetary Economics 49: 265-292.

Tobin, James (1975). "Keynesian Models of Recession and Depression," American Economic Review Papers and Proceedings, May, 65, pg. 195-202.

Woodford, Michael (2003). "Interest and Prices: Foundations of a Theory of Monetary Policy," Princeton University Press.

Woodford, Michael (2009). "Information-Constrained State-Dependent Pricing," Journal of Monetary Economics, forthcoming.

10 Appendix

While the steps in the proofs for propositions 1 and 2 are relatively well known, they are provided here for completeness. For the reader's convenience, I first reproduce a well-known result in linear rational-expectations models from Woodford (2003). I will use his results in my proofs of Propositions 1 and 2 that follow.

10.1 Proposition A.1

Consider a linear rational-expectations model of the form

$$E_t z_{t+1} = A z_t + a e_t,$$

where z_t is a two-by-two vector of nonpredetermined endogenous state variables, e_t is a vector of exogenous disturbances, A is a two-by-two matrix of coefficients, and a is two times the number of exogenous disturbances. Rational-expectations equilibrium is determinate if and only if the matrix A has both eigenvalues outside the unit circle (i.e., with modulus $|\lambda| > 1$). Denote the determinant of the matrix A as $det(A)$ and its trace as $tr(A)$.

The condition for a unique bounded solution is satisfied if and only if

$$\text{Case 1} \quad (a)\ det(A) > 1 \quad (b)\ det(A) - tr(A) > -1 \quad (c)\ det(A) + tr(A) > -1$$

$$\text{Case 2} \quad (d)\ det(A) - tr(A) < -1 \quad (e)\ det(A) + tr(A) < -1$$

Proof: See proof to Proposition C.1 in Woodford 2003, p. 670-71.

10.2 Proof of Proposition 1

The system can be written in the form

$$E_t z_{t+1} = A z_t,$$

where $z_t \equiv \begin{bmatrix} \pi_t \\ \hat{Y}_t \end{bmatrix}$ and $A \equiv \begin{bmatrix} \beta^{-1} & -\beta^{-1}\kappa \\ \sigma(\phi_\pi - \beta^{-1}) & 1 + \sigma\kappa\beta^{-1} \end{bmatrix}$.

Observe that $tr(A) = 1 + \beta^{-1} + \sigma\kappa\beta^{-1}$ and $det(A) = \beta^{-1}(1 + \sigma\kappa\phi_\pi)$. Case 2 of Proposition A.1 clearly does not apply, see, e.g., condition e is violated. Now consider Case 1. As seen above, condition a is satisfied because the determinant is positive. Condition c is also clearly satisfied since both det(A) and tr(A) are positive, which leaves condition b, which we show applies if $\phi_\pi > 1$.

$\beta^{-1}(1 + \sigma\kappa\phi_\pi) - 1 - \beta^{-1} - \sigma\kappa\beta^{-1} > -1 \Leftrightarrow \phi_\pi > 1$, which is what Woodford (2003) calls the "Taylor principle." It is now easy to confirm that the analytic solution in the proposition (which, from the discussion above, we know is unique) is the one given in the proposition, i.e. $\pi_t = \hat{Y}_t = \hat{\Delta}_t = \hat{W}_t = \hat{l}_t = 0$ and $i_t = \bar{r}$. QED

10.3 Proof of Proposition 2

Consider the solution at time $t < \tau$. Given the result from Proposition 1, we can write expectations of inflation and output as

$$E(Y_{t+1} \backslash t < \tau) = (1 - \mu)E_t Y_{t+1}^S + \mu * 0 = (1 - \mu)E_t^S Y_{t+1}^S,$$

where the notation E_t^S is used as the expectation of the variable Y_{t+1} conditional on the shock being in the S state, i.e., $t + 1 < \tau$. Similarly, the notation Y_{t+1}^S is used to signify that this is the value of Y_{t+1} conditional on $t + 1 < \tau$. We can similarly write inflation as

$$E(\pi_{t+1} \backslash t < \tau) = (1 - \mu)E_t \pi_{t+1}^S + \mu * 0 = (1 - \mu)E_t^S \pi_{t+1}^S.$$

Hence

$$Y_t^S = (1 - \mu)E_t Y_{t+1}^S - \sigma i_t^S + \sigma(1 - \mu)E_t \pi_{t+1}^S + \sigma r_S^*$$
$$\pi_t^S = \kappa Y_t^S + \beta(1 - \mu)E_t \pi_{t+1}^S + \delta\psi_S$$
$$i_t = \max(0, r_S^* + \phi_\pi \pi_t).$$

10.3.1 Part i

Consider first part i, and let us conjecture that the zero bound is not binding. Then we can write

$$E_t z_{t+1}^S = A z_t^S + a\psi_S$$

where $z_t^S \equiv \begin{bmatrix} \hat{Y}_t^S \\ \pi_t^S \end{bmatrix}$, $a \equiv \begin{bmatrix} \frac{\delta\sigma\beta^{-1}}{1-\mu} \\ -\frac{\delta\beta^{-1}}{(1-\mu)} \end{bmatrix}$ and $A \equiv \begin{bmatrix} \frac{1+\frac{\sigma\kappa}{\beta}}{1-\mu} & \frac{\sigma(\phi_\pi-\beta^{-1})}{1-\mu} \\ -\frac{\kappa}{\beta(1-\mu)} & \frac{1}{\beta(1-\mu)} \end{bmatrix}$.

Observe that $tr(A) = \frac{1+\beta^{-1}+\sigma\kappa\beta^{-1}}{1-\mu}$ and $det(A) = \frac{\beta^{-1}(1+\sigma\kappa\phi_\pi)}{(1-\mu)^2}$. Case 2 of Proposition A.1 clearly does not apply, e.g., condition e is violated. Hence, consider Case 1. As seen above, condition a is satisfied because the determinant is positive since $\mu < 1$. Condition c is also clearly satisfied since both $det(A)$ and $tr(A)$ are positive. Which leaves condition b, which is shown to apply below:

$$\frac{\beta^{-1}(1+\sigma\kappa\phi_\pi)}{(1-\mu)^2} - \frac{1+\beta^{-1}+\sigma\kappa\beta^{-1}}{1-\mu} > -1 \Leftrightarrow \mu^2 + \mu(\beta^{-1} - 1 + \sigma\kappa\beta^{-1}) + \beta^{-1}\sigma\kappa(\phi_\pi - 1) > 0,$$

which holds for any $\mu \epsilon [0,1]$ as long as $\phi_\pi > 1$. This proves that there is a unique bounded solution in the short run of the form stated in the proposition. The algebraic form of the solution in the proposition can be found using standard methods, e.g., method of undetermined coefficients. Note in particular that the analytical solution for the nominal interest rate suggests that it is only satisfied as long as C1 does not apply. QED.

10.3.2 Part ii

We now consider the case in which both C1 and C2 apply. From the solution for the nominal interest rate in part i of the proposition we see that this implies that the zero bound is binding. Then we can write

$$E_t z_{t+1}^S = A z_t^S + a e_t,$$

where $z_t^S \equiv \begin{bmatrix} \hat{Y}_t^S \\ \pi_t^S \end{bmatrix}$, $a \equiv \begin{bmatrix} -\frac{\frac{\sigma\delta}{\beta}}{1-\mu} & -\frac{\sigma}{1-\mu} \\ -\frac{\delta}{\beta(1-\mu)} & 0 \end{bmatrix}$, $A \equiv \begin{bmatrix} \frac{1+\frac{\sigma\kappa}{\beta}}{1-\mu} & -\frac{\frac{\sigma}{\beta}}{1-\mu} \\ -\frac{\kappa}{\beta(1-\mu)} & \frac{1}{\beta(1-\mu)} \end{bmatrix}$

Observe that $tr(A) = \frac{1+\beta^{-1}+\frac{\sigma\kappa}{\beta}}{1-\mu} = \frac{(1+\beta^{-1}+\frac{\sigma\kappa}{\beta})(1-\mu)}{(1-\mu)^2}$ and $det(A) = \frac{\beta^{-1}}{(1-\mu)^2}$. Case 2 of Proposition A.1 clearly does not apply, because, for example, condition e is violated. Consider Case 1. As seen above, condition a is satisfied because the determinant is positive since $\mu < 1$. Condition c is also satisfied since both $det(A)$ and $tr(A)$ are positive. Which leaves condition b which is shown to apply below:

$$\frac{\beta^{-1}}{(1-\mu)^2} - \frac{(1+\beta^{-1}+\frac{\sigma\kappa}{\beta})(1-\mu)}{(1-\mu)^2} > -1 \Leftrightarrow \beta\mu^2 + \mu(1-\beta+\sigma\kappa) - \sigma\kappa = \mu(1-\beta(1-\mu)) - (1-\mu)\sigma\kappa > 0$$

which is condition C2. This proves that there is a unique bounded solution in the short run of the form stated in the proposition, provided C1 and C2 apply. The algebraic form of the solution in the proposition can be found using standard methods, e.g., method of undetermined coefficients. QED.