ARRA is small, the counterhysteresis effects will be correspondingly small. More important, I think the authors ignore the significant countervailing effect of increased productivity that occurred in this downturn, as firms responded to the slow recovery of demand by laying off workers and changing production methods.

CONCLUSION This paper is important because it develops a benefit-cost framework for evaluating the desirability of fiscal stimulus policy by taking into account the size of the multiplier, the present value of the deadweight costs of future debt service, and the counterhysteresis effects of the improvements in human and physical capital that under certain circumstances accompany the multiplier-generated increase in GDP. The authors are careful to note that their conclusion about the desirability of a fiscal stimulus applies only to the case of a deep recession with monetary policy constrained by the zero lower bound on interest rates. My analysis within their framework, however, does not lead to the same support for the kind of fiscal policy represented by the ARRA. The reason for my skepticism about the positive benefit-cost ratio reflects different evaluations of their three components of the benefit-cost calculation. To summarize:

—I believe the multiplier evidence based on NIPA military spending cannot be applied to budget outlays that do not raise NIPA government spending.

—I think the cost of adding to the deficit includes more than just the deadweight losses associated with the additional future interest costs.

—I view the gains from counterhysteresis effects as overstated both because the multiplier effect is small and because the downturn also induced productivity-enhancing changes in production.

COMMENT BY

VALERIE A. RAMEY This paper proposes the very intriguing idea that government stimulus packages enacted during a severe downturn can be self-financing. Bradford DeLong and Lawrence Summers have actively participated in the public debate on this issue, both in and outside of government, in op-eds and on blogs, so a paper on this topic by them is of particular interest. When I was first asked to discuss the paper, I felt the same kind of anticipation that I did for the 1992 Olympics. Why? Because those were the first Olympics in which athletes who had gone professional were allowed to come back and compete with the amateurs. In the 1992 Summer Games, the U.S. basketball "Dream Team," which included such players as Michael Jordan, Magic Johnson, and Larry Bird, beat every other team

by huge margins and won the gold. However, the returning pros in the subsequent Winter Games, which included such celebrated skaters as Katarina Witt and Brian Boitano, and Jayne Torvill and Christopher Dean, did not do as well. Although they all skated beautifully, the Olympic judges put more emphasis on precision than did the audiences they had performed for as pros. Thus, Torvill and Dean captured only the bronze and the rest did not win medals at all. The question then is, Are DeLong and Summers the "Dream Team," or are they Torvill and Dean?

I will first summarize DeLong and Summers's central idea and highlight some notable elements of their model. Since a key part of their hypothesis is the ability of government spending to reverse hysteresis effects, I will offer some evidence that can be viewed as a test of these effects. I will then assess the notion, which also appears elsewhere in the literature, that the fiscal multiplier may be higher when interest rates are at the zero lower bound and economic slack is high. Finally, I will offer a word of caution on extrapolating from past interest rates.

HYSTERESIS, THE ZERO LOWER BOUND, AND FISCAL STIMULUS DeLong and Summers very clearly present their main idea and their view of how the economy works. They argue that most of the time output is supply-determined and equal to potential output, as in a neoclassical model. Government spending at such times has no impact on output. During times such as the Great Recession and its aftermath, however, output is below potential and is demand-driven. In such a depressed economy, government spending can raise output.

In addition, the authors assume that current output levels can have an effect on future potential output, and thus on actual future output during supply-determined times—a hysteresis effect. Olivier Blanchard and Summers (1986) first introduced the idea of hysteresis effects in the context of lingering high unemployment in Europe, and they appealed to an insider-outsider theory of labor markets to motivate the idea. In the present context, DeLong and Summers appeal to various factors, such as the deterioration of the skills and labor force attachment of the unemployed and the long-term effect on capital of depressed investment rates. They present a very useful summary of the micro evidence on persistence of labor supply decisions as well as macro evidence on the link between financial crises and subsequent output growth.

On top of this structure, the authors also consider the unusually low interest rates prevailing in the U.S. economy today. These low rates potentially have two effects. First, as numerous papers have argued, when nominal interest rates hit their zero lower bound (ZLB), monetary policy becomes

impotent and fiscal policy can become more powerful than usual. Second, the low interest rates make budget deficits very cheap to finance.

This view of how the economy works naturally leads to DeLong and Summers's main conclusion: for a variety of parameter values, short-run increases in government spending during a slump can pay for themselves in the long run. The obvious implication is that fear of large budget deficits should not prevent the federal government from enacting another stimulus package in order to boost aggregate demand. This conclusion is essentially the Keynesian version of so-called supply-side economics, which the Reagan administration used to argue that tax cuts could stimulate the economy enough so that tax revenue would actually rise. Another way to look at it is as the new version of Say's Law: "In a depressed economy, government spending creates its own financing."

IS A SIMPLE MODEL BETTER? DeLong and Summers's model has more in common with standard undergraduate textbook macroeconomic models than with the typical models used at the graduate level or in research papers. For example, there is no discussion of their assumptions about fundamentals such as preferences, technology, and resource constraints. Also, the model is missing the "GE" of DSGE models, for there is no general equilibrium. The paper is silent on why interest rates are so low and why they should be expected to remain so low. Overall, the model is quite stripped down relative to standard modern macroeconomic models.

There are advantages and disadvantages of presenting this idea in such a simplified model. The advantages are several. First, because the model is so stark, the idea is very clearly presented and not obfuscated by inessential technical details. Second, because the idea is not embedded in one of the standard macroeconomic models, it avoids invoking those models' sometimes questionable assumptions.

However, the reason that modern macroeconomics has moved to models with carefully specified assumptions and microfoundations is that without them, one can often end up implicitly making contradictory or dubious assumptions. The reason that one can find many faults with modern macroeconomic models is precisely that they are explicit about their assumptions. Although these models have a long way to go to find better ways to model the economy, I do not think that replacing them with a model based on imprecisely specified intuitive ideas is an improvement. Indeed, I see that as a disadvantage of the approach.

THE KEY RELATIONSHIP IN THE MODEL DeLong and Summers's idea boils down to one key mathematical expression that they derive from their

assumptions about government spending multipliers and hysteresis. According to their model, an increase in government spending to stimulate the economy during a recession can be self-financing (in the sense of generating an increase in annual tax revenue sufficient to pay for the increase in annual debt service) if the following condition holds:

$$r < g + \frac{\eta\mu\tau}{1 - \mu\tau},$$

where r is the interest rate on government bonds, g is the growth rate of potential output, η is the hysteresis effect of current output on future potential output, τ is the tax rate, and μ is the government spending multiplier. The intuition behind this equation is as follows. If output is below potential, it can be spurred by an increase in government spending. The multiplier μ is the measure of how much output rises for a given dollar increase in government spending. This increase in current output is then translated into higher future potential output through the hysteresis effect η. These two effects imply that for plausible values of μ and η and a given tax rate τ, future annual tax revenue will increase more than the required increase in annual debt service. I will devote the rest of my discussion to exploring some evidence on three of the five parameters of this equation: the hysteresis parameter η, the government spending multiplier μ, and the long-term interest rate r.

THE HYSTERESIS PARAMETER DeLong and Summers include a very nice summary of the literature on various features of the economy that could lead to hysteresis. As they acknowledge, however, estimating the extent of hysteresis is very difficult, since it is difficult to distinguish the lingering effects of the recession itself (that is, state dependence) from the continuing effects of the unobserved forces that caused the recession. Nevertheless, the evidence the authors compile is quite interesting. One of the mechanisms they discuss is the potential loss of worker skills or labor force attachment that might result from extended periods of unemployment. The recent evidence presented by Steven Davis and Till von Wachter (2011) of significant and persistent losses in income to workers displaced during a recession is certainly suggestive. As Robert Hall points out in his discussion of that paper, however, whether losses to individual workers represent social losses or just redistributions of rents remains to be seen. If they are the latter, then these losses do not represent actual losses in worker skills.

DeLong and Summers discuss a second mechanism for hysteresis that works through private investment: shortfalls in private investment during a recession, they argue, can lead to persistent effects through reductions in the capital stock. It is certainly the case that a prolonged slump in invest-

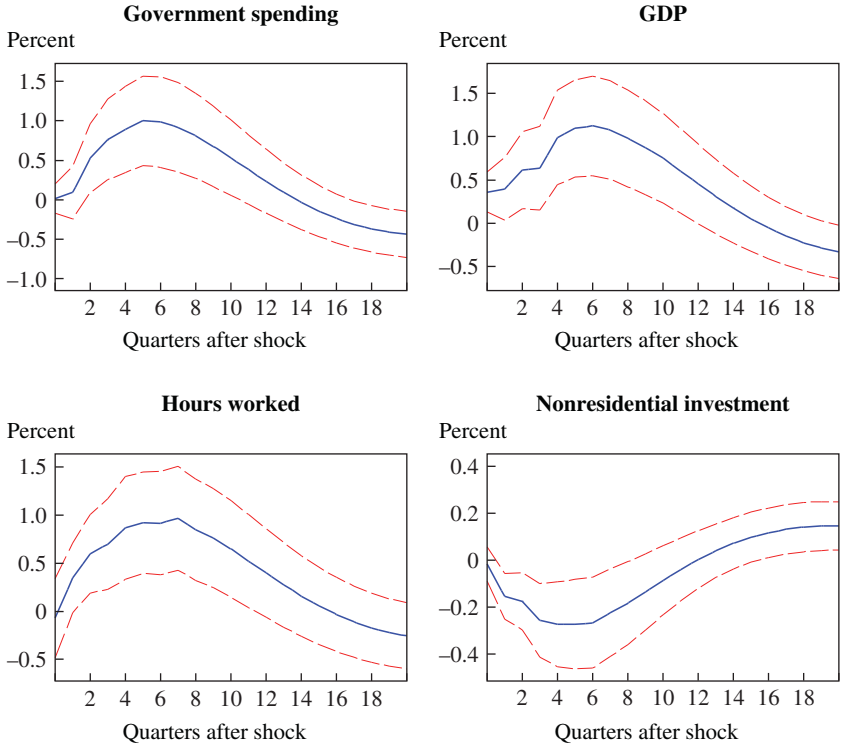
ment can lead to a significantly lower capital stock and hence lower potential GDP. As I will argue below, however, it is not clear that government spending can reverse this effect.

As DeLong and Summers themselves recognize, most of their arguments suggest persistent, but not permanent, effects. A positive depreciation rate on the hysteresis effect can have a sizable effect on their calculations. For example, consider the simulation results in their table 2, which indicate that for η of 0.025 and μ of 1.5, a stimulus package is self-financing as long as the real government interest rate is below 4.95 percent. But suppose, using the same parameters, that the hysteresis effect has a depreciation rate of 10 percent per year. Then only half of the necessary tax revenue is being collected 6 years in the future. Thus, their calculations hinge importantly on their assumption of *permanent* hysteresis effects.

For the sake of argument, suppose that hysteresis effects are indeed permanent. DeLong and Summers's argument still requires another assumption to support their policy prescription of more stimulus spending: they must assume that raising output with government spending can reverse the hysteresis effect. It is not obvious to me that an increase in government spending would create the private investment and skill-building jobs required to do that.

Even without specifying the individual mechanisms, one can test this hypothesis on U.S. data. In particular, if DeLong and Summers are correct that a change in real government spending G raises real GDP Y in the short run, it should have a persistent effect on output. That is, if one can identify exogenous movements in government spending that have led to temporary increases in real output, those increases should have a much more persistent effect on output if there are hysteresis effects. To study this, I use my analysis from Ramey (2011, 2012), which identifies exogenous shocks from military events that generated changes in the expected present discounted value of government spending. I also use a method of identification like that in Blanchard and Roberto Perotti (2002) to show that the results are not limited to just my identification method. The models are estimated from 1939Q1 to 2008Q4. Estimation is by vector autoregressions (VARs) containing log real government spending per capita, log real GDP per capita, the average marginal tax rate from Robert Barro and Charles Redlick (2011), the 3-month Treasury bill rate, log total hours worked per capita, and log real nonresidential investment per capita (all in levels). The Blanchard and Perotti structural VAR (SVAR) identifies the shock to be the shock to government spending, ordered first in the VAR. My EVAR ("expectational VAR") includes my military news variable ordered first and

Figure 1. Impulse Response Functions from an Expectational VAR[a]

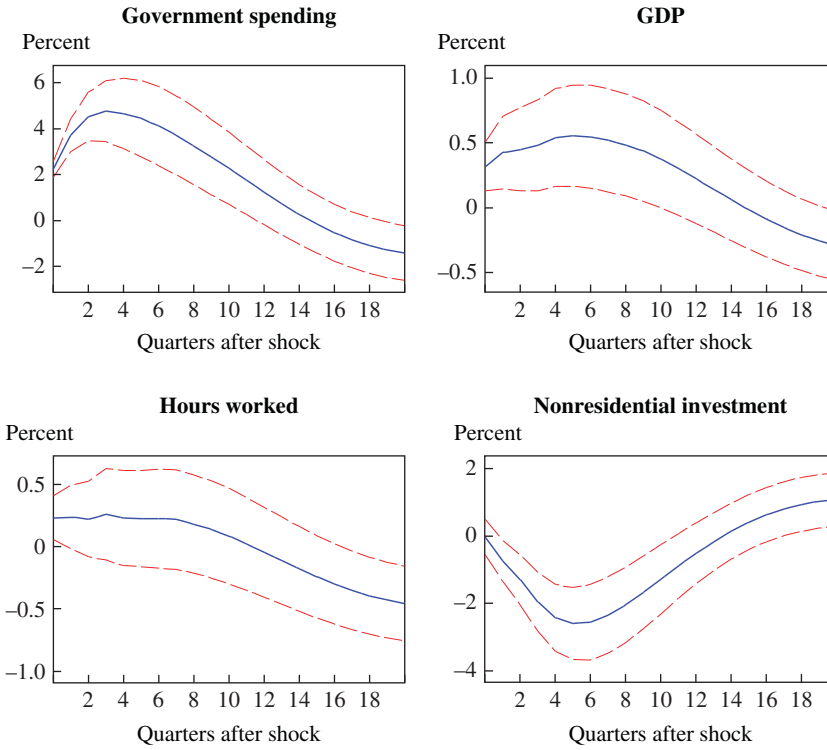


Source: Author's estimates based on the expectational VAR described in the text.
a. The sample is quarterly data for the United States from 1939 through 2008. Dashed lines denote 95 percent confidence intervals based on bootstrap standard errors.

uses shocks to it as the government spending shock. Four lags are used and a quadratic trend is included.

My figures 1 and 2 show the impulse responses (with 95 percent confidence intervals based on bootstrap standard errors) of four of the variables of interest: government spending, real GDP, total hours, and nonresidential investment. In both specifications, a shock raises both government spending and real GDP. They both peak around 6 quarters after the shock and are back to normal by 16 quarters. Total hours also rise, but nonresidential investment falls. Thus, in the historical data, investment is moving in the opposite direction from that which would produce the counterhysteresis effects that the authors argue for. Nor is there evidence of a persistent effect of government spending on real GDP.

Figure 2. Impulse Response Functions from a Structural VAR[a]



Source: Author's estimates based on the structural VAR of Blanchard and Perotti (2002) described in the text.

a. The sample is quarterly data for the United States from 1939 through 2008. Dashed lines denote 95 percent confidence intervals based on bootstrap standard errors.

Since my news variable captures only movements in government spending based on military events, one might wonder whether other types of government spending would have more long-lasting effects. The results using Blanchard and Perotti's framework use a shock to all types of government spending, and yet there is no more evidence of persistent effects on output. In sum, this evidence does not support the notion that an increase in government spending that raises output in the short run has lingering effects on output.

THE GOVERNMENT SPENDING MULTIPLIER A recent paper of mine (Ramey 2012) uses a more precise way to estimate the multiplier in both a VAR and an instrumental variables regression. In particular, it looks at the effect of government spending on private spending ($Y - G$). This method

indicates that the multiplier is significantly below unity—about 0.5 when tax effects are accounted for. DeLong and Summers and numerous others have argued, however, that the multiplier may be higher when there is slack in the economy (Auerbach and Gorodnichenko forthcoming) or when interest rates are at the zero lower bound (Eggertsson 2001).

In principle, it is possible to test this hypothesis on historical data. In work in progress, I have been studying the period from 1933 to 1951. As my figure 3 shows, this period was characterized by very low interest rates, similar to today's, as well as very high unemployment rates for much of the period. Of course, the presence of World War II, with patriotism raising labor force participation rates and controls on the economy dampening consumer spending, make interpretation of the period very complex. It is nonetheless interesting to at least search for differential multipliers during this period.

For this period, I thus estimate the following equation:

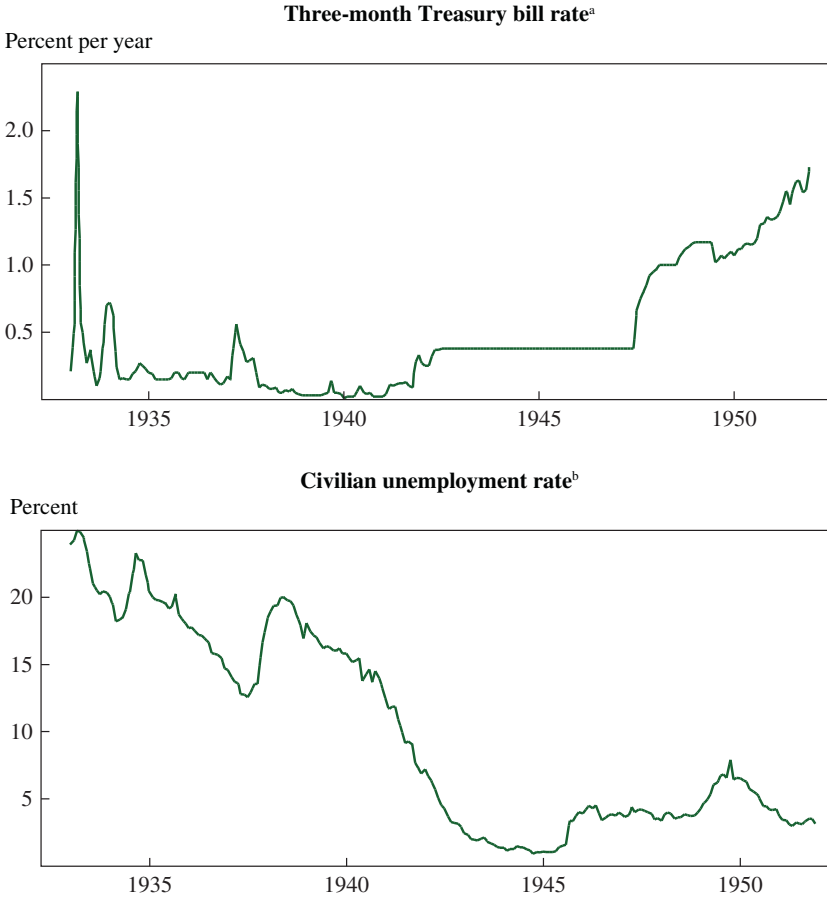
$$\frac{\Delta Y_t}{Y_{t-1}} = \beta_0 + \beta_1 \frac{\Delta G_t}{Y_{t-1}} + \beta_2 \frac{\Delta Y_{t-1}}{Y_{t-2}} + I_t \left[\beta_0' + \beta_1' \frac{\Delta G_t}{Y_{t-1}} + \beta_2' \frac{\Delta Y_{t-1}}{Y_{t-2}} \right] + \varepsilon_t,$$

where I is a dummy variable that takes the value of 1 when the unemployment rate is above 7 percent and zero otherwise, and ε is the error term. I allow all of the coefficients, including the multipliers, to vary according to whether the unemployment rate is above or below 7 percent. The coefficient β_1 gives the multiplier when the unemployment rate is below 7 percent, and $\beta_1 + \beta_1'$ gives the multiplier when the unemployment rate is above 7 percent.

The data used are monthly from January 1933 to March 1951 (the month when the Treasury Accord restoring Federal Reserve independence was signed). The GDP and government spending data are from Robert Gordon and Robert Krenn (2010). The unemployment series are based on my data collection and include emergency workers. To address both the possible endogeneity of government spending and the impact of measurement error from the way that Gordon and Krenn construct their interpolated series, I instrument for government spending growth with lags 2 through 4 of government spending growth (relative to GDP).

My estimate of the multiplier β_1 when the unemployment rate is below 7 percent is 0.581, with a standard error of 0.119. The increment to the multiplier during slack times is estimated to be –0.012, with a standard error of 0.535. Thus, the regression provides no evidence of a higher multiplier during slack times, although the high standard error indicates substantial uncertainty.

Figure 3. Short-Term Interest Rates and the Unemployment Rate, United States, 1933–51



Sources: NBER Macrohistory Database, the Conference Board, *Survey of Current Business,* and author's calculations.

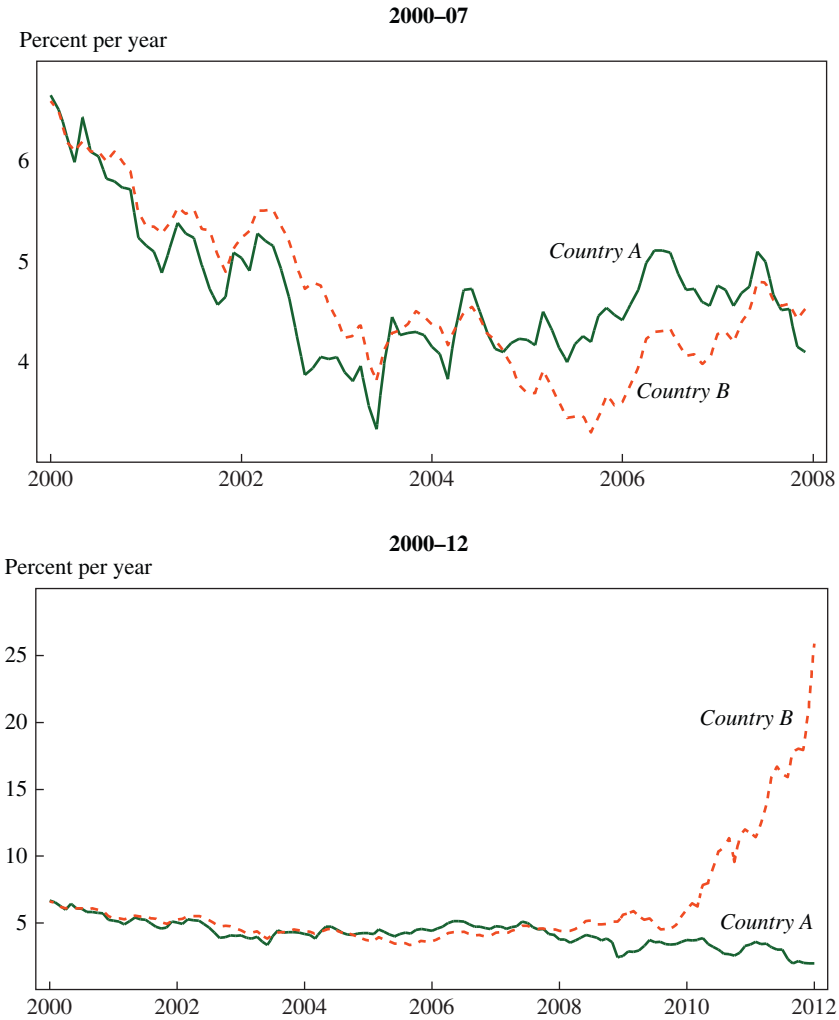
a. NBER Macrohistory Database series 13029b. Data are originally from the Federal Reserve Board.

b. Constructed from monthly series on the numbers of unemployed (including emergency workers) and employed workers. Underlying data from the NBER Macrohistory Database, the Conference Board, and the *Survey of Current Business* were seasonally adjusted, scaled to match the annual unemployment series compiled by David R. Weir, University of Michigan, and then spliced.

This simple analysis finds no support for multipliers that are higher during times of slack and accommodative monetary policies. The analysis presented above is quite simple, and there is some evidence to the contrary in later periods, so more research should be done on this issue.

THE GOVERNMENT INTEREST RATE DeLong and Summers look at historical data on long-term government interest rates to argue that it is unlikely

Figure 4. Long-Term Interest Rates on Government Debt since 2000 in Two Countries



Source: Organisation for Economic Co-operation and Development, Main Economic Indicators.

that interest rates will rise significantly. Here it is wise to bear in mind the Lucas critique (Lucas 1976). In particular, the historical data the authors examine were not generated in a regime in which entitlements were projected to lead to ever-rising deficits as in the current situation.

To illustrate the perils of extrapolating from the past, consider the "tale of two countries" told in my figure 4. The top panel shows interest rates on

long-term government bonds for two countries from 2000 through 2007. Both countries displayed similar patterns in interest rates over this period.

Now consider the paths of interest rates in the two countries when the sample is extended through 2011 (bottom panel). The interest rate for country A remains low while that for country B suddenly explodes and reaches 25 percent. Country A is the United States, and country B is Greece. Simply extrapolating from the past behavior of interest rates would never have led one to predict that interest rates would rise so far in Greece. Thus, current low interest rates should not be taken as a sign of future low interest rates.

CONCLUSIONS DeLong and Summers present the very intriguing idea that government spending can be self-financing when used to stimulate an economy in which output is below potential. Although I have concerns about the lack of rigor of their theoretical model, the idea is still quite interesting. My simple empirical investigations of the hysteresis effect and the government spending multiplier, however, indicate that those two parameters might not be as high as they need to be for this idea to work. Moreover, I have suggested caution in using current low interest rates to forecast the future path of interest rates. Nevertheless, DeLong and Summers have introduced an important new idea that clearly merits future research. This is what we expect from Olympic gold medal winners.

REFERENCES FOR THE RAMEY COMMENT

Auerbach, Alan, and Yurii Gorodnichenko. Forthcoming. "Fiscal Multipliers in Recession and Expansion." In *Fiscal Policy after the Financial Crisis,* edited by Alberto Alesina and Francesco Giavazzi. University of Chicago Press.

Barro, Robert J., and Charles J. Redlick. 2011. "Macroeconomic Effects from Government Purchases and Taxes." *Quarterly Journal of Economics* 126 (February): 51–102.

Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117: 1329–68.

Blanchard, Olivier, and Lawrence Summers. 1986. "Hysteresis and the European Unemployment Problem." *NBER Macroeconomics Annual* 1: 15–90.

Davis, Steven, and Till von Wachter. 2011. "Recessions and the Cost of Job Loss." *BPEA,* no. 2: 1–72.

Eggertsson, Gauti. 2001. "Real Government Spending in a Liquidity Trap." Working paper. Federal Reserve Bank of New York.

Gordon, Robert J., and Robert Krenn. 2010. "The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy." Working Paper no. 16380. Cambridge, Mass.: National Bureau of Economic Research (September).

Hall, Robert E. 2011. Comment on "Recessions and the Cost of Job Loss" by Steven Davis and Till von Wachter. *BPEA* (Fall): 56–61.

Lucas, Robert E. 1976. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

Ramey, Valerie A. 2011. "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics* 126 (February): 1–50.

———. 2012. "Government Spending and Private Activity." Working Paper no. 17787. Cambridge, Mass.: National Bureau of Economic Research.

GENERAL DISCUSSION Robert Hall observed that a better title for the paper would be "Eta," since the paper's surprising results all stem from the authors' beliefs about the value of their hysteresis parameter η. The other parameter values the authors used for their simulations seemed mostly reasonable and uncontroversial to Hall. He noted that although Valerie Ramey had estimated a relatively low value for the multiplier on fiscal spending, the standard error on her estimate was large and did not rule out the possibility that the authors' baseline value of 1.5 was correct. Hall also observed that some alternative ways of analyzing government spending data from World War II generated higher estimates of the multiplier. He found the authors' value for the growth rate reasonable, and although he shared Ramey's concern about the authors' real interest rate assumptions, he thought their baseline value might be reasonable as well.

For the most important parameter, η, however, Hall felt that much more work was needed to arrive at a credible estimate. He noted that for the interesting cases in the authors' analysis, $r - g$ is small, which makes the present value of extra output due to avoided hysteresis significant for decades into the future. In such cases, then, the appropriate value for η would be an average not just over the next decade but over many decades.

Econometrics, however, simply cannot answer the question of whether hysteresis effects, or the effects of avoided hysteresis, are significant far into the future, Hall argued. The "unit root" literature of the late 1980s had found that it was impossible to precisely estimate the persistence of shocks to GDP, yet small differences in the persistence of government spending shocks had very different implications for the analysis. And even if it were possible to estimate the long-run effects of such a shock, the United States has not experienced a government purchases shock in many years. The 2009 stimulus package did not constitute such a shock, as the positive effect of the package on government purchases was slightly more than offset by negative effects from other sources. Hall was skeptical of the suggestion of