The Routes Into and Out of the Zero Lower Bound

Robert E. Hall

The major central banks of advanced countries—the Federal Reserve System, the European Central Bank, the Bank of Japan and those of many smaller countries—are in liquidity traps today, with policy rates at minimum feasible levels. An economy enters a liquidity trap when a shortfall of demand for output calls for a low real interest rate, one so low that, at moderate inflation rates, the Zero Lower Bound (ZLB) on the nominal interest rate imposes a binding lower bound on the real rate. In the United States today, with a policy rate of about 10 basis points and an inflation rate around 180 basis points, the safe short real interest rate is minus 170 basis points, well above the level of around minus 400 basis points that would generate output demand equal to normal levels of output supply.

The basic story is the collision of three forces:

- A decline in output demand—an event without serious consequences in a normal economy,
- The ZLB on the nominal interest rate, and
- Low and stable inflation, so that the implied bound on the real interest rate is constraining.

In the United States and some other countries, notably Spain, the driving force for the decline in output demand was a substantial drop in real-estate values. The drop had a direct effect on household consumption spending—households had been financing a consumption boom with constant increases in indebtedness secured by rising house values. Suddenly, in 2007, the process reversed and households repaid debt, often under compulsion from lenders. The decline in consumption demand, especially spending on consumer durables, began in mid-2007. And, of course, the huge decline in expenditures on homebuilding began at the same time.

The decline in real-estate values also had a large indirect effect through the U.S. financial system. That system has two basic asset classes, real estate and the physical capital of businesses. Households are entirely dependent on financial institutions for real-estate financing. On the other hand, nonfinancial businesses depend mainly on securities markets—stocks, bonds and shorter-term debt securities. Because most household debt is secured by real estate, the result is a financial system highly exposed to real-estate values. The contrast between the 2001 recession and the Great Recession in 2007-09 illustrates the difference. In 2001, the value of business assets, especially tech-related assets, fell dramatically, but the financial system showed no signs of stress. Financial institutions had little exposure to business assets. The stock market communicated losses directly to investors with no banklike intermediation. In the Great Recession, banks and other financial institutions became insolvent or nearly so because of direct and indirect exposure to real-estate values. The stock market fell by about the same percentage in both recessions. In the Great Recession, the fall occurred because the adverse forces from the real-estate crash appeared to threaten a collapse of the whole economy. A large increase in the discount rate applied to risky business returns caused most of the decline in the stock market in 2008 and 2009.

Without the ZLB, the economy would have ridden through the decline in output demand without much damage, because the real interest rate would have fallen enough to keep output near normal. Estimates vary about how negative the rate would have fallen, but the point is that the real rate is the price that clears the output market—

some negative value would do the job. Quite negative real rates in the recoveries from the 1973-75 and 1981-82 deep recessions made possible their V shapes. The nominal interest rate cannot be more than a bit negative, because investors always have the option of holding currency, with a guaranteed save nominal return of zero. With stable low inflation, as the U.S. and all other advanced countries have experienced for a lengthy period, the ZLB on the nominal rate places a bound on the real rate that is a huge constraint on the economy. With, say, 2 percent inflation, the real rate cannot fall below minus 2 percent, which the experience of the past five years teaches is well above the market-clearing rate.

A burst of inflation would permit an adequately low real rate even with the ZLB. The Fed responded aggressively to the events of 2008, ending the year at a zero policy rate. Other central banks followed suit, though not with the same determination. But far from relieving the interest bound, the policy failed to prevent a decline in inflation, a decline that has worsened recently. Fortunately the decline was modest, quite unlike the extreme deflation of 1929 to 1933, which raised the real rate to catastrophic levels.

Though understanding inflation is central to understanding the effect of the ZLB and to the design of countervailing policies, recent experience has shown the defects in economists' earlier thinking about inflation. Since the birth of the Phillips curve in the 1950s, the idea has dazzled macroeconomists that inflation depends on tightness or slack. Yet extreme slack has done little to reduce inflation over the past five years (fortunately!) and extreme tightness in the late 1990s did not result in much inflation.

The obvious conclusion from these observations is that raising prices and wages faster than normal is not a market outcome in a tight economy and raising them slower or even allowing them to fall is not a market outcome in a slack economy. The natural basis for that situation is that markets are in equilibrium, sometimes tight and sometimes slack, but in equilibrium in the sense that no actor believes that changing price- or quantity-related behavior would be privately advantageous.

Because the rate of inflation is completely central to understanding the ZLB and the current and future states of the U.S. and other major economies, I spend some effort in this paper reviewing ideas about equilibriums with variable tightness. All of these ideas rest on a central contribution in macro theory, the Diamond-Mortensen-Pissarides (DMP) model, honored by the Nobel Prize in economics in 2010. The DMP model created a coherent and realistic theory of tightness within which the question can be investigated rigorously. One central implication of the model is that there is no fixed "natural" rate of unemployment which the actual unemployment rate revolves around. Rather, the observed level of unemployment varies according to driving forces and is always an equilibrium.

Since mid-2009, the U.S. economy has expanded slowly and returned partway to more normal conditions, while the economies of the euro area, Britain and Japan have been more stagnant. I discuss the forces that are likely to continue the expansion and ultimately release the U.S. economy from the ZLB, meaning that the economy is back to normal. One is the ebbing of the elevated risk premiums that investors assigned to business income, which held back investment and job creation. Another is the growing shortfall of the capital stock, which has declined since 2008 despite population growth, and is now far below normal, generating a pent-up demand for investment.

I. The Collapse of Output Demand: The Path into the ZLB Economy

In the wake of the financial crisis in September 2008, output and employment fell precipitously in the United States. Although economic activity had begun to decline gradually starting in the previous December, the dramatic decline immediately after the crisis suggests that financial events had a major role in the deep and prolonged slump in the economy. Events in financial markets stand at the forefront of most explanations of the slump.

Commentary has focused on two channels. The first, household deleveraging, emphasizes cutbacks in consumption forced on credit-dependent households by the elimination of opportunities to borrow and by rising requirements to repay existing debt. The second

emphasizes the cutback in plant, equipment and inventory investment resulting from a rising gap between the marginal return to capital and the safe short real interest rate. That wedge rose because financial institutions earned higher spreads between their funding costs and their interest charges, because the institutions rationed credit, and because the equity premium rose substantially. Much of the macroeconomic modeling of the crisis has focused on the first two sources of the wedge, but many sectors of American business have little dependence on debt finance, so the rise in the equity premium is an essential part of the collapse of all forms of investment after the crisis.

In October 2008, the Federal Reserve lowered its policy interest rate to essentially zero, where it remains at this writing. The transition from an earlier policy regime, where the rate responded to current developments, to one that was incapable of further stimulus from lowering the policy rate, was an important feature of the economy in the aftermath of the crisis.

I.i The Financial Wedge

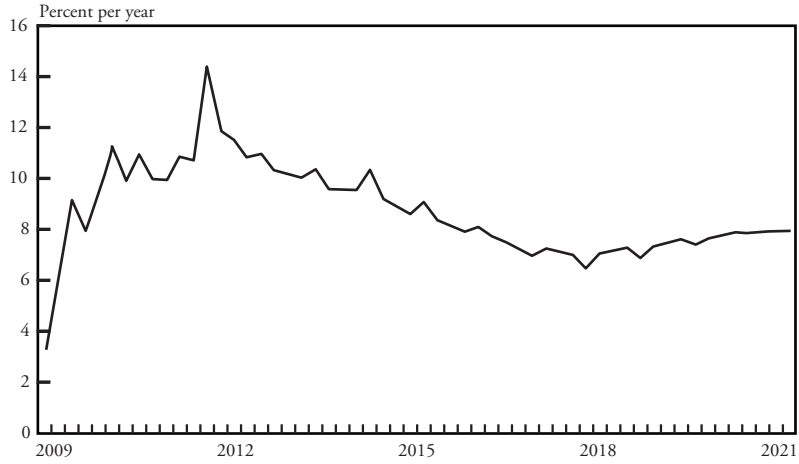
The financial wedge is the difference between the rate of return to capital and the real interest rate:

$$f_t = \frac{1}{q_t}\left[\alpha \frac{y_t}{k_t} + (1-\delta)q_{t+1}\right] - 1 - r_t \qquad (1)$$

Here q_t is Tobin's q, the market value of installed capital, α is the elasticity of output with respect to capital, y_t/k_t is the output capital ratio, δ is the rate of depreciation, and r_t is the safe short real interest rate. The quantity $\alpha \frac{y_t}{k_t}$ is the marginal product of capital with a Cobb-Douglas production function. This calculation is on the same conceptual footing as the investment wedge in Chari, Kehoe and McGrattan (2007), stated as an interest spread.

Chart 1 shows the values of the financial wedge, stated as an annual percent equivalent to a property tax on capital, calculated from Equation 1. The friction began at a low value immediately after the crisis, in the first quarter of 2009, rose to a high level in 2012, then is predicted to decline gradually back to normal over the future. The

Chart 1
The Financial Wedge



Source: U.S. NIPAs including Fixed Asset Accounts.

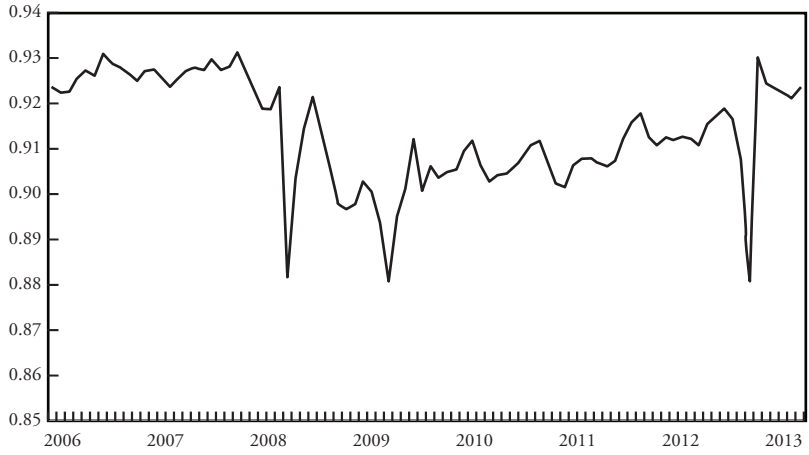
friction is the difference between the quarterly realized return to capital and the risk-free short-term interest rate. There was an immediate decline in investment after the crisis. Tobin's *q* began to fall as investment fell, so the return fell at the same time as the short-term interest rate, and the gap between them—the measured value of the friction—was small. After *q* stabilized, the return to capital was closer to normal, but the short-term rate remained low, so the measured friction was high.

The financial wedge was a leading cause of the depressed levels of output and high levels of unemployment, especially after the middle of 2009.

I.ii Deleveraging

Consumption normally has a stable relation to disposable income. The connection is direct in households with low levels of buffers of liquid assets and little access to borrowing—these households account for somewhat over half of consumption. Among households with scope for smoothing consumption when disposable income falls, the life-cycle-permanent-income model of consumption predicts a rise in the consumption/disposable income ratio if households

Chart 2
The Ratio of Consumption to Disposable Income



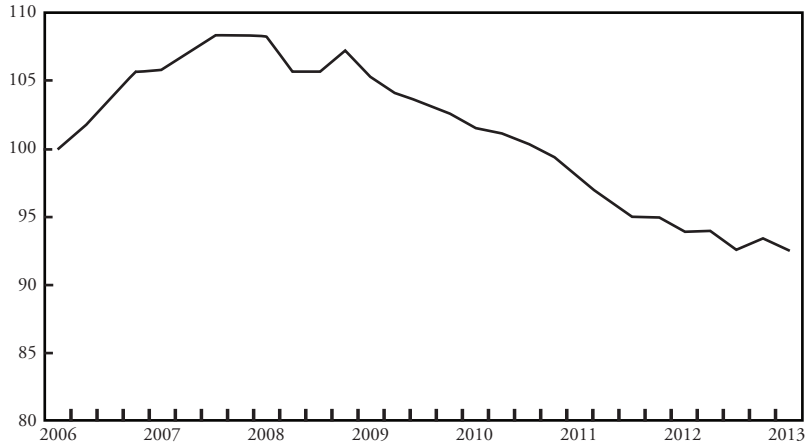
Source: U.S. NIPAs.

perceive a decline in income to be transitory and stability of the ratio if the decline is permanent. Chart 2 shows what happened to the ratio from 2006 to the present.

The ratio plunged starting before the crisis, in the second half of 2007. It reached bottom in early 2009, around the time when real GDP also reached its trough. Then it regained and exceeded its earlier level by the beginning of 2013. Recently it has declined slightly, but is still higher than in 2006. The behavior of the ratio is inconsistent with consumption-smoothing as the dominant response of households to bad times. Rather, it appears that the main force at work was financial stress associated first with the decline in real-estate prices that began in 2007 and the tightening of lending standards that accompanied the price decline. The financial crisis in late 2008 does not appear to have had an important effect on household spending relative to disposable income.

Chart 3 shows that the Federal Reserve Board's Flow of Funds accounts report a large decline in consumer debt—mainly mortgages, car loans and credit card balances—prior to and after the

Chart 3
Real Household Liabilities



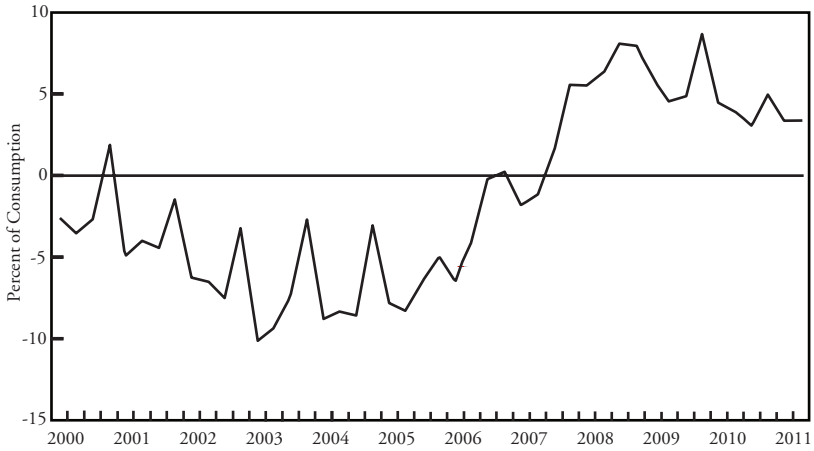
Source: Federal Reserve System Flow of Funds Accounts.

financial crisis, continuing to the present. Households have shed a huge amount of debt in real terms over the past six years.

The decline in debt outstanding is an imperfect measure of deleveraging in the sense of cash flows out of households. Debt will decline with defaults, in which case no corresponding cash outflows squeeze consumption. Notwithstanding the name, the Flow of Funds accounts do not report flows of cash out of households—the flow item for consumer debt is literally the first difference in outstanding debt. The same obstacle to measurement of deleveraging, in the sense that I use the term, applies to any research based on loan balances outstanding.

Chart 4 shows a provisional calculation of the deleveraging flow of cash out of households, adjusted for defaults. The measurement of defaults is a challenge. Banks report a concept called charge-offs to account for losses on loans. The amount represents the bank's estimate of the impairment of the value of nonperforming loans occurring in a given period. The use of charge-offs as an offset to loan balance reductions to measure deleveraging is conceptually appealing, because the charge-off is net of the bank's expected recovery from the sale of the collateral. For example, if a homeowner defaults on a $120,000 mortgage on a house that sells for $100,000,

Chart 4
Burden of Deleveraging as a Percent of Consumption



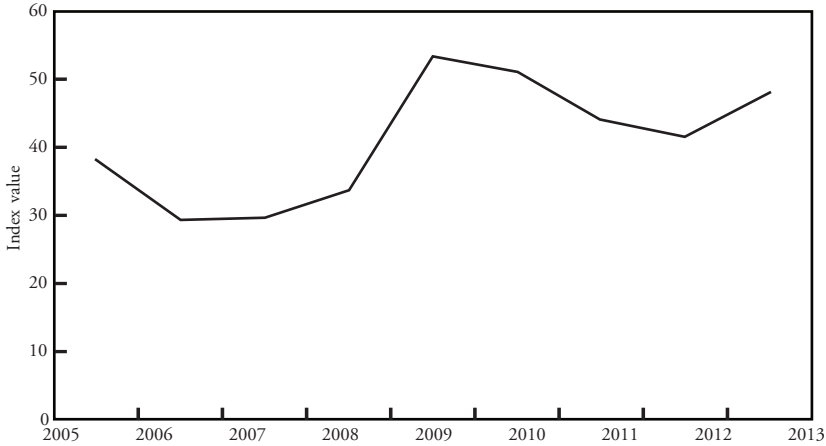
Source: U.S. NIPAs, Flow of Funds, and Federal Reserve data on loan charge-offs.

the charge-off is $20,000. Suppose a new owner buys the house with a no-down-payment loan of $100,000. The household sector has no cash outflow to the financial sector. Outstanding mortgage loans fall by $20,000, the amount of the charge-off. Subtracting the charge-off from the decline in outstandings gives the right answer of cash outflow from households of zero. In the case of unsecured credit-card lending, it is immediately apparent that cash outflows to lenders from households is net of charge-offs.

Relying on banks' estimates of charge-offs may distort the timing of estimated cash outflows from deleveraging. Saulny (2012) reports that it is common for banks to leave defaulted homeowners in their homes to act as caretakers. Whether banks report full charge-offs for houses in this situation is not known—the low market value of banks with large mortgage portfolios relative to the book values of those portfolios suggest that there may be lags in updating book values. The book value of a loan declines each time a charge-off is reported on the loan.

Chart 5 shows that household financial stress continued long after the financial crisis abated. It shows an index of Google searches for "withdrawal penalty." The substantial increase in concerns about withdrawal penalties from retirement plans and longer-term savings

Chart 5
Google Searches for "Withdrawal Penalty"



Source: Google.com/trends.

instruments confirms a persistent increase in household willingness to incur penalties to prevent deep cuts in living standards.

II. Framework for Understanding the Effect of the Collapse of Product Demand

The real interest rate clears the current output market. Investment spending—plant and equipment, homebuilding and consumer durables—falls with higher real rates. Net exports fall. Even consumer spending on nondurable goods and services falls, because a higher real rate encourages saving and deferral of spending. A higher real rate may also affect the supply of output, though I will defer this topic and treat supply as inelastic for the moment. Figure 1 shows the basic idea about the joint determination of the real interest rate and output that lies at the heart of every modern macroeconomic model.

The chart shows output demand under normal conditions as a solid line. The corresponding real interest rate on the vertical axis is at its normal level. The chart also shows a second output demand function in a dashed line where demand is lower at every real interest rate, thanks to the negative shock. The effect is to lower the real interest rate while leaving the level of output unchanged. The real rate is a complete shock absorber. Note that the decline in the real rate may

involve the central bank, which lowers its policy rate as soon as it detects the decline in demand. Obviously I am omitting dynamics here.

Now suppose that, for some reason, the real rate is bounded; it may not drop below a stated level. Figure 2 shows an economy subject to a negative demand shock twice as large as in Figure 1, where the shock drives down the equilibrium real rate below the bound. The bounded rate, shown as a horizontal line, intersects demand at a point where it is below supply. The economy has an excess supply output on account of the bound.

II.i Inflation and the Real Rate of Interest

Before pursuing the central question of how the economy deals with the excess supply, I need to say more about the connection between a bound on the real rate and the bound that is causing so much trouble today around the world, which is on the nominal interest rate. By definition, the real rate is the nominal rate less the rate of inflation. The discussion now enters the imperfectly understood realm of the determination of the rate of inflation.

The bound on the nominal interest rate is at zero. The reason is that a nominal rate below zero would make currency a dominant way to hold wealth, because its nominal return is zero. If a central bank tried to enforce a significantly negative nominal rate for a nontransitory period, investors would demand large volumes of currency, in preference to securities with negative returns. The nominal interest rate is bounded by the negative of the cost of safe storage of currency, which I will take to be zero, though in fact it is slightly positive. Thus the ZLB on the nominal rate implies that the real rate cannot be lower than minus the inflation rate. Binding lower bounds are a disease limited to economies with low inflation rates, a point emphasized by a number of prominent macroeconomists recently. We made the ZLB crisis by adopting what seemed to be healthy policies of low inflation.

An economy with a completely flexible price level could find a standard equilibrium after a deep negative shock to output demand. To

Figure 1
In Equilibrium, the Real Interest Rate is at the Level that Equates Output Demand to Supply

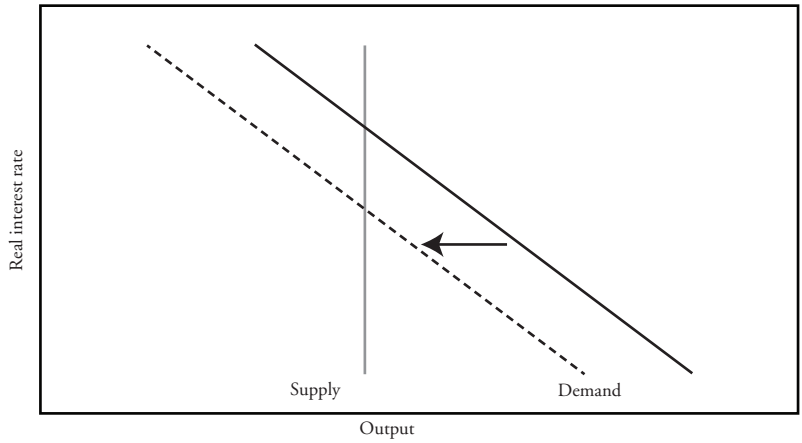
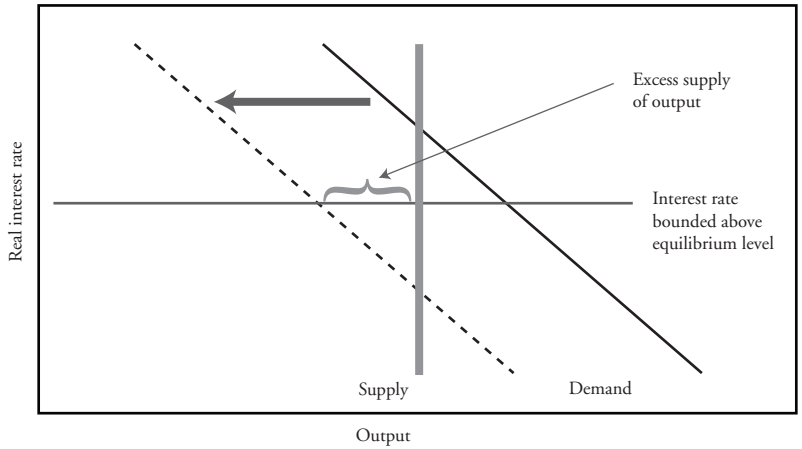


Figure 2
Excess Supply of Output when the ZLB Binds



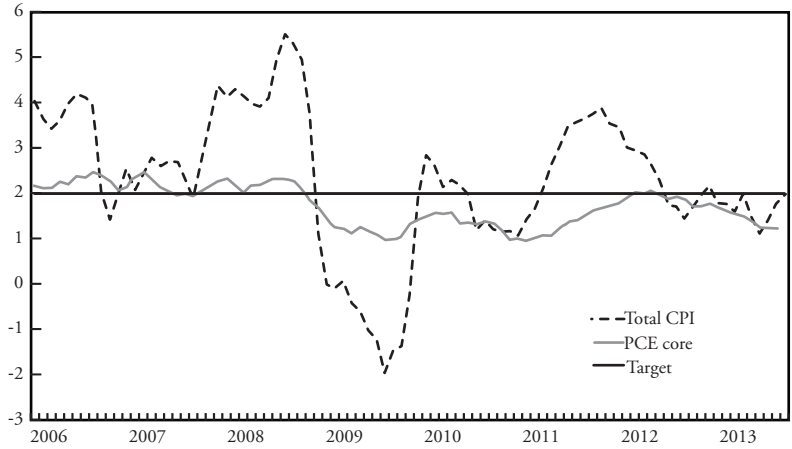
achieve the needed low real interest rate, the rate of inflation needs to be sufficiently great to keep the nominal rate above zero despite the low (probably negative) real rate needed to maintain output at the level of supply.

Achieving high enough rates of inflation has eluded central bankers in most countries subject to the ZLB. The Taylor rule provides a good framework for studying this issue. Broadly speaking, the Taylor rule calls for expansion when a weighted average of (1) the inflation rate less its target and (2) output less its target, is negative. All major central banks of advanced countries have been deep in the territory where their Taylor rules call for expansion, but have been unable to deliver that expansion, as year after year has passed with shortfalls from their targets. The Fed, in particular, has seen more than four years with inflation and output both below target.

The Great Depression brought vicious deflation, resulting in a compounding of the initial adverse shock, as real rates rose to extreme levels. More recently, Japan has suffered from high real rates during the long slump that began more than two decades ago. Apart from Japan, modern advanced countries now at the ZLB have generally seen only small declines in inflation. One of them, Great Britain, somehow achieved an increase in inflation despite substantial slack.

Prior to the recent deep worldwide recession, macroeconomists of all schools took a negative relation between slack and declining inflation as an axiom. Few seem to have awakened to the recent experience as a contradiction to the axiom. In a Jackson Hole paper three years ago, Stock and Watson (2010) showed not only that inflation was essentially unresponsive to the deep slump following the Great Recession, but that earlier evidence suggesting such a relationship in U.S. data was misunderstood. The historical pattern is that a rise in unemployment generates a transitory decline in inflation, but the rise wears off quite quickly, and an extended period of high unemployment—as in the U.S. since 2007—has no effect on inflation. Important support for this proposition is available in the opposite direction, from the experience of the late 1990s, when unemployment fell to extremely low levels without an outbreak of inflation.

Chart 6
Two Measures of U.S. Inflation



Source: U.S. NIPAs and Bureau of Labor Statistics.

Chart 6 shows the U.S. inflation rate from 2006 to the present. The chart has two measures, both stated as 12-month percent changes. The first is the standard Consumer Price Index of the Bureau of Labor Statistics, often called "headline inflation." The other is the core inflation index the Fed generally favors, the monthly price index from the National Income and Product Accounts for consumption apart from food and energy. The two indexes tell the same story; though, as intended, the core measure has less volatility. The horizontal line shows the Fed's target, 2 percent. Prior to 2008, inflation was close to target. A burst of commodity price increases drove inflation upward in 2008 and then in 2009, those prices fell back to normal, depressing overall inflation. At the same time, the sharp contraction associated with the financial crisis lowered inflation as well. Inflation fluctuated above and below target from 2010 through 2012. The continuing drop so far in 2013 has stirred some faint concerns that the deflation feared in 2009 may finally be striking in 2013 or 2014.

Chart 7 shows that wage inflation tells an even simpler story over the same period. It charts the Employment Cost Index (including fringe benefits), an index that adjusts carefully for the composition of the labor force. The Fed has no target for wage inflation, though some economists believe that it would be a better target than price

Chart 7
U.S. Wage Inflation

Source: Bureau of Labor Statistics, Employment Cost Index.

inflation. In the longer run, wage inflation should exceed price inflation by the rate of growth of productivity. Taking that as about 1 percent per year yields a target of 3 percent to correspond to the price inflation target of 2 percent. The chart shows that wage inflation ran below target at all times from 2008 to the present. Wage inflation fell a little over 1.5 percentage points at the outset of the crisis, then recovered by more than a percentage point while unemployment was in the 9 to 10 percent range, and finally has fallen recently by about half a percentage point. Again, there is faint concern about a potential move to deflation in the wage data. Note that the overall movements of wage inflation show no support whatever for the hypothesis that slack brings a persistent and growing decline in inflation. Rather, as Stock and Watson found, inflation drops at the outset of a contraction and then returns to normal level when the economy reaches bottom, with extensive slack and high unemployment.

There is a widespread belief among macroeconomists that the stability of inflation at positive, though slightly diminished rates, in the face of high levels of unemployment, reflects strongly anchored expectations about inflation. Certainly this idea could help explain the difference between the favorable experience following the crisis of 2008, when real rates remained at reasonable levels, and the disaster

following the crisis of 1929, when rampant deflation drove real rates to high levels. The public became adapted to stable low inflation in the decades before 2008, whereas inflation was highly unstable in the two decades before 1929. But the model has yet to appear that embodies anchoring in a persuasive way. The New Keynesian model has a major role for expectations, but only because of a mechanical feature preventing businesses from responding to shifts in demand except with a substantial lag.

A fair statement about macroeconomic understanding of inflation today is that inflation is highly persistent and resistant to any force, policy or otherwise, that might change it. As Michael Woodford explained in his Jackson Hole paper a year ago, the New Keynesian model has the property that an announcement by a central bank that it planned to inflate aggressively as soon as it emerges from the ZLB will speed the exit from the ZLB. Given the shaky foundations of the model and the low likelihood that a central bank would actually follow through on the policy when the time comes, it is understandable that central bankers have been reluctant to provide this kind of forward guidance. The Fed has limited its guidance to a firm statement that, as long as inflation is below target and unemployment is well above its normal level, it makes sense to continue aggressive expansionary policies.

II.ii Reasons Why Inflation Does Not Respond to Slack or Tightness

The traditional thinking behind inflation determination as described in the Phillips curve and formalized in the New Keynesian model is that the economy generally has a gap between supply and demand and inflation falls when the gap is positive and rises when it is negative. With excess supply, sellers undercut each other gradually to try to take profitable business away from their rivals. Inflation is lower than normal while this process occurs.

A logical explanation for the lack of such a response to low output is that sellers do not have profitable opportunities to take business away from rivals by cutting prices. Instead, they are in some kind of equilibrium where their current prices are optimal and it is optimal to change the prices at the established prevailing rate.

This possibility leads me to investigate the question: In what sense might the economy be in high-unemployment equilibrium during a lengthy period when the lower bound is binding, such as 2009 to the present? My discussion is within the only fully-developed modern theory of unemployment, that derived from the work of DMP. The investigation has two benefits. First, to the extent that high unemployment is an equilibrium, the stability of inflation in the presence of persistent high unemployment is less of a mystery. The tradition of regarding high unemployment as a disequilibrium that gradually rectifies itself by price-wage adjustment may rest on a misunderstanding of the mechanism of high unemployment. Second, a better understanding of high unemployment may aid the development of policies that could speed the return of unemployment to normal levels.

II.iii The DMP Model

I will begin the discussion with a sketch of the basic principles of the DMP model that will operate in all variants of the model that I consider later. The canonical expression of the model is Mortensen and Pissarides (1994). Shimer (2005) is a more recent derivation that stays close to the data for the U.S. economy and Hall (2009) is a detailed statement of the way that I express the model here.

The DMP model focuses on the job-creation decision of the employer. When an employer adds a worker, the employer gains the present value of the difference between the worker's marginal contribution to revenue (the marginal revenue product of labor) and the worker's pay. This present value is the job value. But to reach the point where this gain occurs, the employer expends recruiting effort. The net benefit to the employer is the job value less the cost of recruiting a worker. With free entry to hiring, employers push recruiting effort to the point where the net benefit is zero. Thus the job value controls the amount of recruiting effort.

The second key element of the DMP model is a positive relation between recruiting effort and the speed with which job seekers find jobs. When employers are making high effort—posting many vacancies and advertising their existence—job seekers succeed in finding jobs quickly. Unemployment is then low. Recruiting effort determines the tightness of the labor market.

Putting these two principles together leads to the conclusion that the job value determines unemployment—a high job value results in low unemployment. The last issue is what determines the job value. Here is where the current crop of variants of the DMP model differs among themselves. In the canonical DMP model, the worker and employer make a wage bargain according to the Nash bargaining principle. That version also assumed, implicitly, that the product market was competitive, so the marginal revenue product is the marginal product of labor. Shimer (2005) showed that the job value is unlikely to change much under those assumptions, because the wage tracks the marginal product closely, leaving the difference between the two essentially constant. He also observed that the marginal product itself has little cyclical movement. Shimer set off a scramble to find reasonable alterations in the job-value part of the DMP model that generate larger variations in the job value and hence explain the large movements of unemployment. Some of the resulting proposals are helpful in understanding what happens in an economy when it hits the lower bound on the interest rate.

I consider four variants of the DMP model that deliver high unemployment in an economy where the lower bound on the interest rate interferes with the equilibrium shown in Figure 1.

II.iv Mortensen

Two years ago I wrote a paper that pointed out a conflict between the standard analysis of the rise in unemployment that occurs when an economy hits the lower interest-rate bound and the DMP model. The standard analysis treats unemployment as a variable that can rise when the economy is in the situation described in Figure 2. When the demand for output falls short of the supply, enough workers lose their jobs to lower supply to the level of demand. My point was that models invoking that principle failed to specify a version of the theory of unemployment founded in DMP principles that was consistent with the large increase in unemployment. Nothing transmitted the shortfall of demand in the output market to the labor market. In particular, no mechanism was present in those models to depress the job value.

Mortensen (2011) is a comment on that paper that offers a simple solution to my challenge. His model takes the rate of inflation as fully predetermined, so the binding ZLB on the nominal interest rate fixes the real interest rate at minus that rate of inflation. Accordingly, the demand for output is below the normal level. The job value is below normal, so unemployment is higher than normal—enough higher so that lower employment accounts for the lower level of output. In effect, the model shifts output supply to the left, so that it intersects the demand function at the lower level of output where the fixed real interest rate cuts the demand function, as shown in Figure 3.

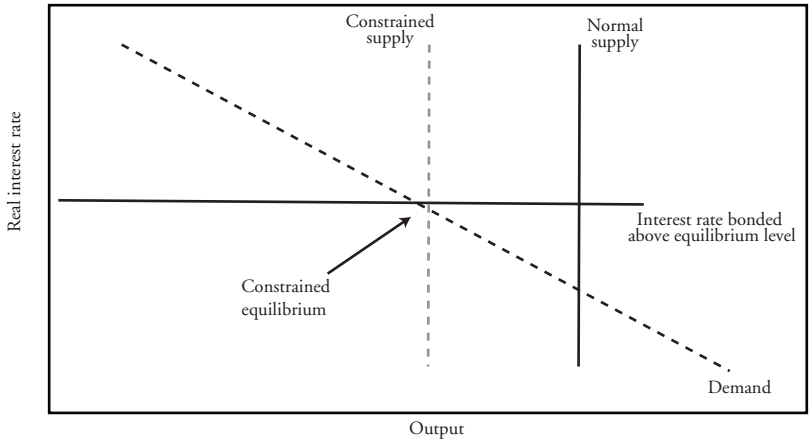
To generate the lower job value, Mortensen supposes that there is a final output industry that does not hire labor from a DMP-style labor market. Rather, it purchases intermediate inputs from a variety of differentiated sellers. Its output is constrained by the binding ZLB, which has the effect of lowering the prices it pays for its inputs. The intermediate-product makers hire workers in a DMP market. The job value they receive is lower than normal because of the depression in the value of their output. Thus, according to standard DMP principles, their job values are lower, they recruit less intensively, and the unemployment rate is higher. Equilibrium in the model occurs where unemployment is sufficiently high that the combined output of the intermediate sector is low enough so that the final goods producers only produce the level of output demanded at the interest rate dictated by the ZLB.

The Mortensen model does not describe a full equilibrium, though many macroeconomists may find its constrained equilibrium descriptive of conditions in an economy depressed on account of the ZLB. Final producers could make money by purchasing more inputs from the intermediate sector and offering purchasers a slightly better deal than is prevailing in the output market. That market is uncompetitive in an unexplained way.

II.v Walsh

Walsh (2003), writing well before the ZLB was a concern outside of Japan, and thus not focusing on the ZLB issue, proposed a way to integrate the DMP model into the New Keynesian sticky-price

Figure 3
Mortensen's Model of Low Output when the ZLB Binds



framework. Employers in his model have market power, so the variable that measures the total payoff to employment is the marginal revenue product of labor in place of the marginal product of labor in the original DMP model. Price stickiness results in variations in market power because sellers cannot raise their prices when an expansive force raises their costs, so the price-cost margin shrinks. Rotemberg and Woodford (1999) give a definitive discussion of the mechanism, but see Nekarda and Ramey (2010) for negative empirical evidence on the cyclical behavior of margins. Hall (2013b) discusses this issue further with additional negative evidence. The version of the New Keynesian model emphasizing price stickiness suffers from its weak theoretical foundations and has also come into question because empirical research on individual prices reveal more complicated patterns with more frequent price changes than the model implies.

Walsh adopts the Nash wage bargain of the canonical DMP model, which implies that his model may generate low unemployment responses for the reason that Shimer (2005) pointed out. Conceptually, it remains the case that Walsh was the first to resolve the clash between Keynesian models with excess product supply and the DMP model of unemployment.

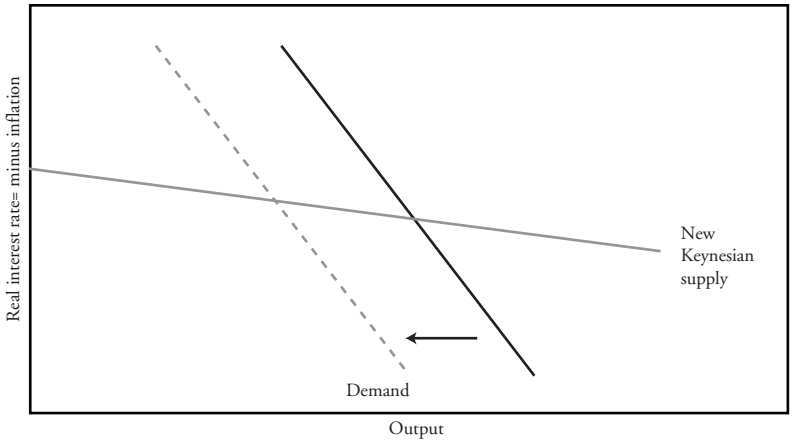
Walsh's model has a sticky output price, not a fixed one. The rate of inflation responds to the level of output in accord with the New Keynesian Phillips curve. Consequently, the real interest rate is an active variable. At the ZLB, the real rate is minus the rate of inflation. Figure 4 shows the effect of a downward shift in output demand. The rate of inflation falls. The marginal revenue product of labor falls, so the job value falls, unemployment rises and output falls. With a sufficiently flat supply function, the outcome matches the observed events in the U.S. following the financial crisis—inflation fell a bit and unemployment rose substantially. Note that this chart refers to the part of the decline in demand that occurs after the economy hits the ZLB. The vertical axis refers to both the real interest rate and the rate of inflation, which are the negatives of each other because the nominal rate is pinned at zero. The part of the decline in demand that occurs before hitting the ZLB causes a decline in the real rate, but this can't be shown in a two-dimensional graph.

Implicit in this view is a possibility of a highly unstable economy, as DeLong and Summers (1986) pointed out. If inflation is more responsive to slack, so that the supply function has close to the same slope as the demand function, a small shift in demand could generate a huge change in output.

II.vi Gertler-Sala-Trigari

A second New Keynesian proposal—more widely accepted currently—introduces a nominal element into wage determination. The canon of the modern New Keynesian model, Christiano, Eichenbaum and Evans (2005) has workers setting wages that are fixed in nominal terms until a Poisson event occurs, mirroring price setting in older versions of the New Keynesian model. That paper does not have a DMP labor market. Gertler, Sala and Trigari (2008) (GST) embed a DMP labor-market model in a general-equilibrium model, overcoming Shimer's finding by replacing Nash bargaining at the time of hire with a form of wage stickiness. Gertler and Trigari (2009) developed the labor-market specification. A Poisson event controls firm-level wage bargaining, which takes the Nash form. Between bargaining times, the wage of newly hired workers adheres to the most recent bargain. If labor demand turns out to be higher than

Figure 4
Output and Inflation Declines from Negative Shock to Output Demand in an Economy with a New Keynesian Phillips Curve



expected at bargaining time, the part of the surplus captured by the employer rises and the incentive to recruit workers rises. By standard DMP principles, the labor market tightens and unemployment falls. Though the model is Keynesian in the sense of sticky wages, it describes an equilibrium in the labor market in the sense of Hall (2005)—the relation between workers and an employer is privately efficient. GST build a model of the general-equilibrium response to monetary and other shocks in a version of the Gertler-Trigari setup where the wage bargain is made in nominal terms. The GST paper resolves the clash by making the DMP determination of unemployment sensitive to the rate of inflation. It does not treat the ZLB on the nominal interest rate explicitly, though it contains all the elements necessary for that analysis.

A key idea in Gertler and Trigari (2009), put to work in the GST paper, is that workers hired between bargaining times inherit their wage terms from the most recent bargain. In principle, this setup could violate the private efficiency criterion by setting the wage too high to deliver a positive job value to the employer or too low to deliver a job value below the job candidate's reservation level, but,

again, in practice this is not likely to occur. If it were an issue, the introduction of state-dependent bargaining would solve the problem, at the cost of a more complicated model.

The GST model assumes that the wage bargain is made in money terms, as the traditional Keynesian literature likes to say. The substance of the assumption is that a state variable—the most recently bargained nominal wage—influences the job value for new hires until the next bargain occurs. This assumption has had a behavioral tinge in that literature—the role of the stale nominal wage arises from stubbornness of workers or employers or from money illusion. From the perspective of bargaining theory, however, as long as the stale wage keeps the job value in the bargaining set, that wage is an eligible bargain. See Hall (2005) for further discussion, not specifically in the context of a nominal state variable. There's no departure from strict rationality in the GST model.

The implications of a model linking the current job value to a stale nominal variable are immediate: The more the price level rises from bargaining time to the present, the higher is the job value in real terms. A sticky nominal wage links inflation and unemployment in the way required by Figure 4.

II.vii Hall

In Hall (2013a), I consider a source of fluctuations in the job value that has been implicit in the DMP model from the outset, but has escaped attention, so far as I know. The job value is the present discounted value of the future difference between a worker's productivity and the worker's pay. Even if that difference is unaffected by a negative shock, if an increase in discount rates accompanies the shock, the job value will decline and unemployment will rise accordingly.

Of course, a crisis results in lower discounts for safe flows—the yield on five-year U.S. Treasury notes fell essentially to zero soon after the crisis of late 2008. The logic pursued in my paper is that the flow of benefits from a newly hired worker has financial risk comparable to corporate earnings, so the dramatic widening of the equity

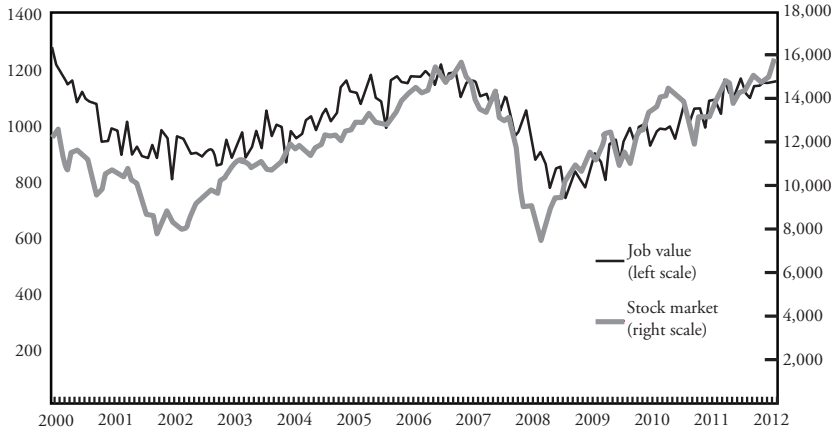
premium that occurred in the crisis implied higher discounting of benefit flows from workers at the same time that safe flows from Treasuries received lower discounting. In the crisis, investors tried to shift toward safe returns, resulting in lower equity prices from higher discounts and higher Treasury prices from lower discounts.

My paper does not explain why risky flows receive higher discounts in recessions. Rather, it documents that fact by extracting the discounts implicit in the actual stock market. I use the framework of modern finance theory, where the discounter is stochastic and present values are the expected products of the discounter and the stochastic future cash flow. Then I demonstrate the plausibility of the hypothesis that the same stochastic discounter also applies to the net benefit of hiring a new worker.

The basic proposition that the stock market varies largely because of changes in discount rates is the conclusion of a famous paper, Campbell and Shiller (1988). Cochrane (2011) discusses the finding extensively. When the value of the stock market is high, relative to dividends or earnings, the discount rate is low and present values are high. In bad times—recessions and early in recoveries—the market is low, the discount rate is high and present values are low. The discount rate is closely related, but not identical to, the expected return on a stock-market investment. In good times, real interest rates are normal or above (2 percent per year or higher), whereas expected real returns in the stock market are below normal (6 percent or lower). Thus the equity premium—the difference between the expected return in the stock market and the interest rate—is low in good times. In really bad times the nominal interest rate is at its lower bound of zero, the real interest rate is, say, minus 2 percent, while the real return in the stock market may be 15 percent. The equity premium is high, sufficiently high that the discount rate is abnormally high even though the real interest rate is low. A central feature of a realistic macro model that comprehends the ZLB is a highly variable equity premium.

How does the ZLB on the safe nominal rate affect the discount on future business that rises in a crisis? Consider the case where a force that boosts the nominal interest rate raises the real rate by the same amount (because it leaves the rate of inflation unchanged). In

Chart 8
Job Value from JOLTS Compared to Wilshire Stock-Market Index



Source: Job Openings and Labor Turnover Survey, Bureau of Labor Statistics, and FRED.

a simple but arguably realistic analysis, the force leaves the equity premium unchanged, so the expected return in the stock market is elevated by the same amount that the lower bound has elevated the safe nominal interest rate.

Chart 8 shows the job value calculated from data in the Job Openings and Labor Turnover Survey (JOLTS), together with the Wilshire index of the broad stock market, deflated by the price index for GDP and detrended. The Wilshire includes almost all of the value of publicly traded U.S. corporations. The similarity of the two series is remarkable. The chart strongly confirms the hypothesis that, however asset markets value uncertain future payoffs, the valuation of the total payoff to corporations and the valuation of the payoff to employers from their workers results in quite similar movements of the resulting values.

The evidence that the job value moves along with the stock market has two implications for output in the post-crisis economy and in other contractions. First, events that trigger a rise in financial discounts, such as a financial crisis, will lower job values substantially, causing a corresponding increase in unemployment and decline in output. In other words, there is a direct linkage from financial disturbances, not

just a response operating through a decline in the demand for output. A financial crisis has a direct adverse effect on output supply through the discount channel. Second, other forces connected to a financial crisis, such as a decline in real-estate prices, cause declines in output demand. The ZLB may block a decline in the discount rate that would have had a favorable effect on the job value and thus cut unemployment and increased product supply. As noted earlier, absent the ZLB, the discount rate would have fallen. That fall would have offset part or all of the increase in unemployment.

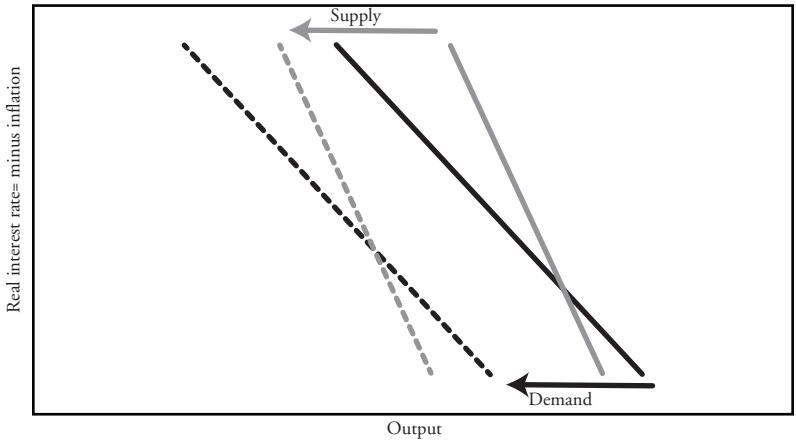
II.viii Conclusions About Equilibrium With High Unemployment

Among the modifications of the DMP model that may aid understanding of high unemployment and stable inflation at the ZLB, I believe that GST's comes closest to meeting the challenge of describing an equilibrium. The stability of inflation arises from the inertia in the wage bargain. That bargain delivers a wage that lies in the bargaining set of a worker and employer, so it is an equilibrium in the usual sense that no bilateral rearrangement can make the two parties better off.

My paper on the widening of the risk premium that accompanies a dramatic contraction of output demand also describes an equilibrium, but its role is somewhat different than is GST's, because it could explain the observed decline in output only be accident. A given rise in discounts raises unemployment by a corresponding amount, which is not necessarily equal to the observed amount. Unlike GST, where the rate of inflation adjusts to equate supply and demand, my model has no free variable to perform that role.

Figure 5 illustrates how the GST and discount channels could account for what happened in the U.S. and other economies after they hit the ZLB. Both the demand and supply functions shifted to the left—demand for the reasons discussed at the beginning of the paper and supply because the same forces also caused a big increase in the risk premium and hence depressed the job value, raised unemployment, and cut output. The role of the GST channel appears as downward slope of supply, but the supply function is not nearly as flat as

Figure 5
ZLB Analysis with Shifts in Both Demand and Supply



in Figure 4, so it cuts demand in the opposite direction. Output falls a lot and inflation falls a little, in line with what actually happened.

The traditional Phillips curve view, built into Figure 4, is that inflation is extremely resistant to slack—it takes very high unemployment to lower inflation significantly. An account of the post-crisis economy that relies entirely on the GST channel has that uncomfortable property. On the other hand, in Figure 5, inflation is quite sensitive to slack. The reason that inflation falls only slightly is that the discount channel shifts supply down by about the same amount as the shift in demand, so only a small decrease in inflation is needed to bring them into balance.

Another interesting feature of Figure 5 is that it accounts for the variety across countries in inflation and output responses. Britain experienced a significant increase in inflation after the crisis, along with a major decline in output. In the chart, this outcome corresponds to a smaller downward supply shift relative to the decline in demand.

III. Monetary Policy

It's fairly obvious that monetary policy does not have instruments to restore ZLB economies to their normal conditions; else much more progress back to normal would have occurred. The Fed has

undertaken a huge expansion of its portfolio and announced that it will continue to keep the funds rate close to zero and maintain an expansionary stance until inflation breaks out or unemployment approaches normal. This combination has not yet closed much of the shortfall in output. Unemployment remains well above any reasonable target and inflation below a reasonable target and forecasts are for a continuation of those clear signs of inadequate stimulus. Both quantitative easing and forward guidance, as implemented by the Fed, are obviously weak instruments.

III.i Woodford's Analysis of Monetary Policy

Woodford (2012) is a wonderfully detailed review of the evidence on forward guidance and quantitative easing. His important conclusions are (1) there is a good case to reduce the 25-basis-points rate the Fed currently pays on reserves, (2) forward guidance needs to take the form of commitments to expansionary policies, rather than mere forecasts, and (3) expanding the Fed's portfolio is not in itself very effective—targeted asset purchases offer the best hope of expansionary effect.

With respect to the interest paid on reserves, there seems to be a general failure to appreciate that paying an above-market rate on reserves changes the sign of the effect of a portfolio expansion. Under the traditional policy of paying well below market rates on reserves, banks treated excess reserves as hot potatoes. Every economic principles book describes how, when banks collectively hold excess reserves, the banks expand the economy by lending them out. The process stops only when the demand for deposits rises to the point that the excess reserves become required reserves and banks are in equilibrium. That process remains at the heart of our explanation of the primary channel of expansionary monetary policy. With an interest rate on reserves above the market rate, the process operates in the opposite direction: Banks prefer to hold reserves over other assets, risk adjusted. They protect their reserve holdings rather than trying to foist them on other banks. An expansion of reserves contracts the economy. The Fed could halt this drag on the economy by cutting the rate paid on reserves to zero or perhaps minus 25 basis points.

The only excuse for not cutting the reserve rate is the belief that short rates would fall and money-market funds would go out of business. This amounts to an accusation that the funds are not smart enough to figure out how to charge their customers for their services. Traditionally, funds imposed charges in a range of 4-50 basis points, in the form of deductions from interest paid. A money-market fund using a floating net asset value can simply impose a modest fee, as do conventional stock and bond funds. The SEC may accelerate this move by requiring all money funds to use floating NAVs.

Forward guidance needs to take the form of a credible commitment now to expand more in the future than purely forward-looking policies would call for. The central bank has to promise to deviate in the expansionary direction from its hard-earned reputation for having solved the fundamental commitment problem of avoiding the temptation to overexpand. That's hard to accomplish.

Woodford's case for commitment to monetary-policy rules is persuasive. He observes that price-level targets have an advantage over inflation targets because they build in expansion after a ZLB period, with substandard inflation, as the economy has above-normal inflation to regain its committed price-level trajectory. I do not share Woodford's enthusiasm for nominal GDP targeting, for the reason in Hall and Mankiw (1994): The volatility of productivity growth is quite high. A policy of stabilizing nominal GDP growth would require contractionary policies to lower inflation when productivity growth is unusually high. Such a policy might easily trigger a spell at the ZLB.

III.ii Multiple Equilibriums

The Taylor rule has become the standard way to think about monetary rules. It provides an indirect nominal anchor by specifying a feedback rule from the inflation rate to the nominal interest rate. Woodford (2003) remains a definitive statement about monetary policy in this environment, notwithstanding the absence of "money" from its title.

Whether a Taylor rule pins down the price level is the subject of a rich literature. Almost all of the analysis is within the New Keynesian

framework. Those of us with longtime reservations about that framework—reinforced recently by the failure of substantial slack to bring declining inflation—do not have a substitute analysis of interest-rate feedback rules. We do have an understanding of other policies, such as the gold standard, but no inclination to recommend their adoption.

Until recently, policies that appeared to follow Taylor rules performed well in practice, so the theoretical possibility of indeterminacy seemed academic. The ZLB effectively suspends the Taylor rule—if the rule calls for a nominal interest rate of zero or higher, the ZLB is not binding. I am not aware of any aspect of post-crisis experience that raises a suspicion of indeterminacy, but I respect the importance of the issue and hope advances in analysis occur, especially in company with a model of inflation that comes closer to explaining the puzzle of its lack of a relation to slack.

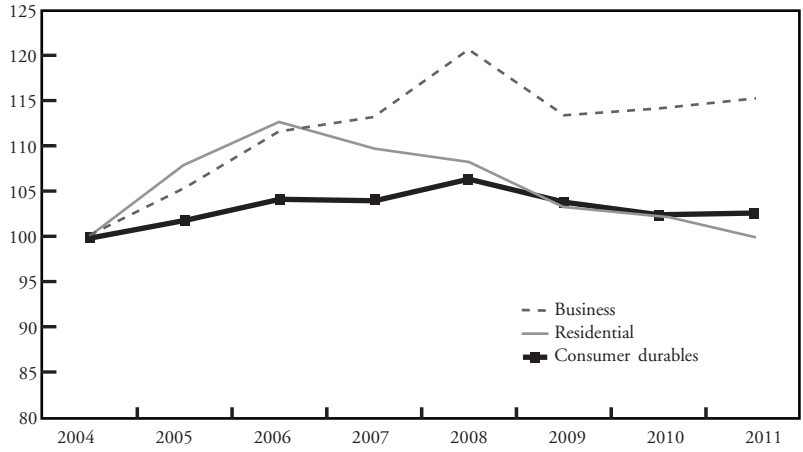
IV. The Route Out of the Zero Lower Bound

Most of the developments that led the U.S. and other advanced countries into ZLB slumps are self-correcting. In the U.S., some evidence suggests that deleveraging pressure on households has subsided. The substantial rise in the stock market since 2009 means that the risk premium for business income is more or less back to normal. Investment flows are beginning to return to normal. In the labor market, the job value is already back to normal, but unemployment is still well above normal. The major potential exception to the good news is the hint of a move toward deflation.

IV.i Financial Wedge and Investment

Chart 1 shows that the financial wedge is well below its peak value and is expected to continue declining to normal over the next few years. The rise in the stock market since its trough in 2009 coincides with a return of the risk premium—a key element of the wedge—to normal levels already. The extra inhibition on plant and equipment investment from the wedge is largely eliminated already, so as output continues to recover, investment should return to normal.

Chart 9
Stocks of Business, Residential and Consumer Physical Capital



Source: Fixed Asset Accounts, U.S. NIPAs.

Investment in all three forms—business plant, equipment and software; homebuilding; and consumer durables—has been far below normal since 2007. As a result, capital stocks have fallen in real terms, as Chart 9 shows, far below their normal growth paths. There is a pent-up demand for investment in the three categories, as existing stocks have aged and investment has fallen far short of deterioration. Standard investment theory projects rising investment levels from depleted stocks even without other improvements in the economy.

IV.ii Household Deleveraging

The recovery of the ratio of consumption to disposable income shown earlier in Chart 2 is a good sign that deleveraging is no longer squeezing household spending as hard. The dramatic decline in real household liabilities shown in Chart 3 confirms that financial burdens on households are well below crisis levels. On the other hand, the frequency of searches for "withdrawal penalty" on Google ticked upward this year and is not too far below its crisis value.

IV.iii The Labor Market

Chart 8 shows that the labor market is back to normal in terms of the job value—employers have the same incentive to create jobs as they did in 2006. In the JOLTS, the source for the data on the job

value, the time required to fill a vacancy (the ratio of the stock of vacancies to the flow of new hires) is at its normal level, after falling to a much lower level at the trough in 2009. On the other hand, the unemployment rate, at 7.6 percent in May 2013, is about 2 percentage points above normal, indicating continuing slackness in the labor market. The disagreement between the employer's perspective—a labor market back to normal—and the job seeker's perspective—jobs still quite hard to find—is the subject of a recent large and interesting literature, surveyed in Hall (2013a). At least part of the explanation is that a much higher than normal fraction of the unemployed today became unemployed through permanent job loss, rather than by quitting or being laid off with a possibility of rehire. At all times, that category those who lost a job has lower job-finding rates. Similarly, many who have lost a job have reached unusually high durations of unemployment, where again job-finding rates are lower. Over the coming couple of years, the labor market should work off the backlog of hard-to-place unemployed.

Another important factor in the labor market is the unusually low level of labor-force participation. Slack conditions in the labor market resulted in withdrawal from the market. Historical experience suggests that the participation rate will return to normal, but more slowly than the return of unemployment to normal. The decline in participation adds to the adverse shift of supply that occurs when unemployment rises, but is corrected more slowly.

IV.iv The Deflation Nightmare

So far, inflation has fallen only slightly and remains in positive territory. Fears in early 2009 that rapid deflation might break out and cause the economy to collapse as in 1929 to 1933 proved unfounded, luckily. I have advanced the hypothesis that rampant price-cutting has failed to appear because businesses are in equilibrium and perceive that price-cutting has bigger costs than benefits. If the hypothesis is wrong and businesses are finally responding to five years of slack by cutting prices, the generally optimistic tone of this section could be quite mistaken. The bottom could fall out of the economy as it did in the Great Depression.

V. Concluding Remarks

The central danger in the next two years is that the Fed will yield to the intensifying pressure to raise interest rates and contract its portfolio well before the economy is back to normal. The worst step the Fed could take would be to raise the interest rate it pays on reserves. The analysis of this paper focusing on the zero lower bound applies equally to a reserve rate above zero. Every percentage point increase in the reserve rate drives the real interest rate up and contracts the economy by the principles discussed here.

With respect to policies that might lower the probability of a repetition of the multitrillion dollar disaster of the past five years, it is true that a policy of higher chronic inflation would have given monetary policy more headroom for expansion to counteract the decline in output demand and to prevent it from causing a decline in output. But I see that response as distinctly second-best. Much preferable are policies to maintain a robust financial system that responds smoothly to declines in real-estate prices. Requiring more capital in financial institutions is an important part of good policy, but to determine the amount of capital, there is no substitute in a modern financial system for frequent and rigorous stress-testing. Derivatives create exposures that are not recorded as leverage, but are fully apparent in stress tests. With a stable, bullet-proof financial system, policies of low inflation are quite safe.

References

Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1 (3), pp. 195-228.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2007. "Business Cycle Accounting,"*Econometrica*, 75 (3), pp. 781-836.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113 (1), pp. 1-45.

Cochrane, John H. 2011. "Presidential Address: Discount Rates," *Journal of Finance*, 66 (4), pp. 1047-1108.

DeLong, J. Bradford, and Lawrence H. Summers. 1986. "Is Increased Price Flexibility Stabilizing?," *American Economic Review*, 76 (5), pp. 1031-1044.

Gertler, Mark, and Antonella Trigari. 2009. "Unemployment Fluctuations with Staggered Nash Wage Bargaining," *The Journal of Political Economy*, 117 (1), pp. 38-86.

_____, Luca Sala, and Antonella Trigari. 2008. "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining," *Journal of Money, Credit and Banking*, 40 (8), pp. 1713-1764.

Hall, Robert E. 2013. "High Discounts and High Unemployment," Hoover Institution, Stanford University.

_____. 2013. "What the Cyclical Response of Advertising Reveals about Markups and other Macroeconomic Wedges," Hoover Institution, Stanford University.

_____. 2009. "Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor," *Journal of Political Economy*, 117 (2), pp. 281-323.

_____. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness," *American Economic Review*, 95 (1), pp. 50-65.

_____, and N. Gregory Mankiw. 1994. "Nominal Income Targeting," in N. Gregory Mankiw (ed.), *Monetary Policy*, University of Chicago Press, pp. 71-93.

Mortensen, Dale T. 2011. "Comments on Hall's Clashing Theories of Unemployment," Department of Economics, Northwestern University.

_____, and Christopher Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, pp. 397-415.

Nekarda, Christopher J., and Valerie A. Ramey. 2010. "The Cyclical Behavior of the Price-Cost Markup," University of California, San Diego.

Rotemberg, Julio J., and Michael Woodford. 1999. "The Cyclical Behavior of Prices and Costs," in John Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics*, Volume 1B, Amsterdam: North-Holland, Chapter 16, pp. 1051-1135.

Saulny, Susan. 2012. "When Living in Limbo Avoids Living on the Street," *New York Times*, March 4.

Shimer, Robert. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95 (1), pp. 24-49.

Stock, James H., and Mark W. Watson. 2010. "Modeling Inflation After the Crisis," National Bureau of Economic Research, working paper 16488.

Walsh, Carl E. 2003. "Labor Market Search and Monetary Shocks," in S. Altug, J. Chadha, and C. Nolan (eds.), *Elements of Dynamic Macroeconomic Analysis*, Cambridge University Press, pp. 451-486.

Woodford, Michael. 2012. "Methods of Policy Accommodation at the Interest-Rate Lower Bound," Columbia University.

_____. 2003. *Interest and Prices*, Princeton University Press, Princeton, New Jersey.